Science-Metrix

Network and bibliometric mapping of climate change adaptation and resilience research

Date:
October 29, 2021

Submitted to:
Bruce Currie-Alder
International Development Reseach Centre
Ottawa, Canada

By:

Science-Metrix Inc.
1335 Mont-Royal E. ▪ Montréal ▪ Québec ▪ Canada ▪ H2J 1Y6
1.514.495.6505 ▪ 1.800.994.4761
info@science-metrix.com ▪ www.science-metrix.com

Contents

Tables

Figures

Acronyms

AUID	Scopus author unique identifier
CCARD	Research on climate change adaptation and resilience for development
FCDO	Foreign, Commonwealth & Development Office
HIP	Percent share of publications in a set falling amongst the top 10% or 30% of publications from their subfield for their interdisciplinarity integration score
HMP	Percent share of publications in a set falling amongst the top 10% or 30% of publications from their subfield for their multidisciplinarity integration score
IDRC	International Development Research Centre
II	Interdisciplinary integration index
LMIC	Lower-middle income countries, as defined by the Organisation for Economic Co-operation and Development Official development assistance list
LIC	Low income countries, as defined by the OECD Official development assistance list
MI	Multidisciplinary integration index
OA	Open access
OECD ODA	Organisation for Economic Co-operation and Development Official Development assistance list
SDG	Sustainable development goals of the United Nations
UMIC	Upper-middle income countries, as defined by the Organisation for Economic Co-operation and Development Official development assistance list
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization

Executive summary

Study background and objectives

The International Development Research Centre (IDRC) is a major funder of climate change adaptation and resilience research on the global stage. The IDRC climate change team is also involved in international governance processes through its participation in intergovernmental fora such as the UN Environment Programme. To support its strategic planning activities, the IDRC climate change team sought bibliometric mappings of recent developments in the field of interest. Specifically, IDRC holds an ongoing interest in measuring the extent to which recent efforts have been cross-disciplinary, published as open science, and engaged in North–South collaborations. It also wishes to determine whether cross-disciplinary, open access and collaborative North-South research has changed network topology in the field, and supported the advancement of women and South-based authors in particular. In progressing this agenda, IDRC collaborates with the UK Foreign, Commonwealth & Development Office (FCDO).

To provide IDRC with some insight on these issues, Science-Metrix produced a network analysis of co-authorship characteristics in scientific publishing in the field of climate change adaptation and resilience research.

The questions that guided this work are:

- What are the shape and composition of research networks in the field of climate change adaptation and resilience?
- Are southern actors and/or women actors becoming more central over time? Have some of these actors become key connectors?
- Have[1] multi-disciplinary and multi-country collaborations impacted network topologies and especially positions for southern and/or women actors?

Answers to these questions were provided by conducting an author-level network analysis of between 40,000 to 750,000 global authors, and up to 385,000 publications (depending on the analysis) in the field of climate change adaptation and resilience. Networks were constructed using co-authorship of publications as recorded in Scopus (with co-publication links as edges and individual authors as nodes). Q1 was answered by computing network heterophily (e.g. diversity in collaboration patterns across the North-South and gender dimensions) and descriptive statistics on the characteristics of publications in various breakdowns of the field. Q2 was answered by computing author-level average centrality of women and southern authors relative to men and northern authors. Q3 was answered by recomputing these women and southern authors average centralities in breakdowns of the field aligned with the explanatory variables of cross-disciplinarity practice, North-South co-publications and open access publishing. Additional descriptive statistics were computed on IDRC-FCDO publications to provide some preliminary insights into related program outcomes. Note that the main network of climate change adaptation and resilience research was defined as a combination of UN Sustainable Development Goal (SDG)-aligned research (i.e. research falling within a subset of 4 SDGs). To be included in the analysis,

[1] This report does not explore how collaboration impacts network typology

research publications had to have a clear regional engagement with southern localities. That is, only publications with the mention of a city, region or country located in the Global South were included in the resulting dataset on climate change adaptation and resilience research for development (CCARD).

Specification of the study's scope

Network analysis is an established technique for analysing research portfolios and is increasingly used in formal program evaluations and outcomes assessments. In support of the scope of this study, it was decided, at inception, to focus on the field of climate change adaptation and resilience as a whole, rather than try to specifically capture the outcomes of IDRC- (and its partner FCDO) funded consortia, or other North-South collaboration projects and consortia. The design was informed by a rationale to capture general field-level dynamics that influence author-level centrality, rather than to harvest program outcomes.

Working with a large dataset of authors and publications inevitably introduces some boundaries to the study's scope. For instance, because of the expert manual curation that would have been required, it was not possible to determine how participation in specific research fronts, or affiliation with specific research institutions, were associated with increasing or higher author-level centrality for women or southern researchers. The study design did not allow for the identification of specific countries where author-level or women author-level centrality changed in recent years. Yet it was possible to measure author centralities within the aggregate of researchers from low-income countries, lower-middle income countries, and those participating in North-South collaborations.

Limitations

Within the study scope defined above, there are some limitations to our findings that require an initial explanation. The study compared centralities for southern and women authors, and heterophily in the main study data set against the same measures in "subnetworks" of cross-disciplinary publications and North-South co-publications. These comparisons suggest possible trends in the fields, yet to isolate explanatory variables would have required regression model. Deploying such models to the numerous combinations of outcome and explanatory variables was beyond the scope of this study. To explore this limitation further, Science-Metrix has volunteered to conduct one such regression analysis for IDRC, as a cost-neutral, additional work step. The regression analysis will focus on one outcome variable (most likely author-level centrality) using a restricted set of explanatory variables (see section below on "Follow-up regression analysis work").

Key findings from the study

(1) Over the analysis period, women have increased their relative centrality (while remaining at a gap to men) and slightly increased their participation in authorship. Women's average centrality relative to men has moved from -38% to -29% between 2010-2014 and 2015-2019. They account for 33% of publication authorships (defined as author-publication pair, with a single author or publication being associated with multiple authorships) in the most recent period, against 31% and 29% in the two prior periods. Authors have a slight tendency to collaborate with authors of the same gender, and this has been stable over time (heterophily measured at around 0.90).

(2) Men tended to have high levels of average centrality even in those publications with high levels of authorship by women. Even in the small subset of publications (less than 10% of the field) where women provided 75% or more or authorships, contributing men tended to make up a comparatively high proportion of the network's most central authors (6 or 7% in the two recent periods, against 11-12% for women).

(3) Southern authors have a lower average centrality compared to Northern authors, and the gap has increased over time. Despite the design decision to focus on publications with clear regional engagement in Global South countries, South authors appear to be behind North authors in terms of their relative centrality in the field's network. The gap in centrality has actually increased over time (from -20% to -24%, and then to -33%). While the network saw heterophily on the North-South dimension, there has been a tendency towards more homophily in time (heterophily moving from 1.23 to 1.14 and then 1.08). That is, North-South co-authorships tend to represent a diminishing share of the overall amount of authorship relationships in the field.

(4) North-South co-publications have been associated with a decrease rather than an increase in South author centrality, and the resulting gap has increased over time (from -34% to -40% and then -44%; note that this trend mirrors a similar trend in the main dataset)**.** Considering only publications with at least one contribution by a researcher based in a low-income country, South authors leads were decreasing in time (difference in average centrality for South authors went up from +58% to +70% between 2005–2009 and 2010–2014, but went down to do +29% in 2015–2019). In lower-middle income country-publications, the difference in average centrality for South authors against North authors decreased from +20% to +11%, and then almost reached parity at +3% in the latest period.

(5) The position of women relative to men has improved over time across all sets of cross-disciplinary publications considered, although the difference in average centrality remains negative. For instance, in highly multidisciplinary publications, those negative differences have reduced from –33% (two earlier periods) to -17%. Women held roughly as many authorships in the subsets of highly multidisciplinary and interdisciplinary publications as in the field overall.

(6) On balance, cross-disciplinary research likely had a negligeable impact on South author positions and potentially a negative impact on one dimension. Highly multidisciplinary publications tended to include more North-South co-publications (recording a share of 45% such collaborations, against 32% in the overall publications). Yet, in the 2010–2014 and 2015–2019 periods, South authors' centrality within highly cross-disciplinary publications followed closely the levels and trends found in the overall publication set. For example, the difference of South author centrality against North author centrality increased from -25% to -33% over those periods. The exception was for the most highly interdisciplinary research $HIP_{10\%}$, where South authors recorded a much larger gap in average centrality than North authors compared with the corresponding gaps recorded in the overall publication set.

(7) North-South co-publications likely did not meaningfully impact positions for women in the network. North-South co-publications recorded slightly less women authorships than in the overall field (31% to 33% in the 2015–2019 period). Differences in average centrality between women and men did not differ meaningfully in those publications. The gap in average centrality for women in North-South co-publications has decreased by roughly the same magnitude as in the main dataset (from -40% to -29%

between 2010–2014 and 2015–2019). Publications with a high proportion of women authorships (75% or more) recorded much lower values for international co-publications (18% against 38% in the main dataset) and for North-South co-publications (15% against 32%). This observation might provide a signal (by no means definitive) that women authors are less likely to collaborate internationally; or, alternatively, that the available international collaborators tend to be men rather than women, creating a potential trade-off between collaborating with other women or collaborating internationally.

(8) Differences on the centrality of women within Open Access publications compared to other publications were both positive and negative, depending on the individual indicator. Women have made gains in average centrality in Open Access (OA) publications over time (from a difference of-38% in 2010–2014 to -25% in 2015–2019), replicating the trends found in the main dataset and in non-OA publications. OA publications did see higher women average centrality than non-OA publications in the recent period, but this pattern was not stable in time.

(9) Open Access publishing may be associated with lower levels of average centrality for South authors, although this dimension did not appear to affect temporal trends in South author centrality. OA publications on climate adaptation and resilience recorded higher levels of international co-publications (46%) and North-South co-publications (32%) than their non-OA counterparts. Yet, OA publications saw a much wider gap for South authors in centrality (-38%) than non-OA publications (-22%) in the recent period. OA publication replicated the trend already found in the overall dataset of decreasing relative average centrality for South authors.

(10) IDRC-FCDO articles published between 2015 and 2019 contained highly collaborative, cross-disciplinary and open research. 60% of the 403 publications included in the sample were North-South co-publications. Almost 30% of these publications were either amongst the most highly multidisciplinary and/or interdisciplinary in their fields (or put differently, three times the reference figure of 10%). Two-thirds of these publications were available under various OA modalities. Yet IDRC-FCDO funding did not appear to have increased shares of authorships by women: one-third of IDRC-FCDO publications' authorships were held by women, very close to the overall figure found in the field under study. It would require a more sophisticated *evaluation* design (as opposed to the network analysis conducted here) to evaluate robustly the specific outcomes enabled by IDRC-FCDO.

Preliminary design for a follow-up regression analysis work

The results obtained so far provide initial signals on how multi-country collaboration, multidisciplinarity and Open Access practices have supported Southern and women author centrality. Science-Metrix proposes to conduct further work in the form of a regression analysis focusing on one main outcome variable of interest and a small selection of explanatory variables, based on the signals already obtained. We suggest that the regression analysis should include one or more independent variable(s) that are direct targets of intervention in IDRC-FCDO programs. For example, probing the influence of "North-South co-publications" (core explanatory variable) on "South author centrality" or "women author centrality" (outcome variables). To take another example, if OA publishing is enforced by IDRC for its grantees, but IDRC does not control the diversity in disciplinarity backgrounds of its funded teams, then probing the "influence of OA publications on South author centrality" may have more program-level implications than probing the "influence of multidisciplinary publications on South author centrality". Below is a high-

level design for a potential regression analysis. As this is illustrative, Science-Metrix cannot guarantee at this stage that any of the variables presented will make final inclusion in the model:

Outcome variable:

- Average centrality: author-level, based on co-publication links

Core model specification:

- Model for southern authors
- Model for women authors

Main explanatory variables

- Share of co-authors in author' publications that are women (author-level women-men heterophily)
- Share of author's publications that are women-men co-publications
- Share of co-authors in author' publications that are southern (author-level southern-northern heterophily)
- Share of author's publications that are North-South collaborations
- Overall interdisciplinarity score of papers by the author
- Overall multidisciplinarity score of papers by the author

Additional explanatory variables that can include other dimensions that were of interest in the present study, or confounding factors that need to be tested for:

- Share of publications as first, senior or corresponding author
- Share of authors' publications supported by IDRC-FCDO
- Share of IDRC-FCDO's total publications to which the author has contributed
- Share of OA publications by the author
- Number of publishing years (as a proxy for seniority)
- Number of articles by the author in the CCARD subfield
- Number of articles by the author in all subfields
- Main country of author
- Single most common Topic of prominence of authors' publications

Contribution from each of these variables to the variance on author-level centrality can be computed, with a view to determine which variables explain most of the differences observed in a given model.

1 Introduction

1.1 Context

Canada's International Development Research Centre (IDRC) is a major funder of climate change adaptation and resilience research on the global stage. IDRC is also involved in international governance processes through its participation in intergovernmental fora such as the UN Environment Programme. To support program design, IDRC sought bibliometric mappings of recent developments in the field of interest. It has an ongoing interest in measuring the extent to which recent efforts have been cross-disciplinary, published as open science, and engaged in North–South collaborations. It is also interested in determining whether cross-disciplinary, open access and collaborative North-South research has changed network topology in the field and supported the advancement of women and southern authors in particular. Science-Metrix produced a network analysis of co-authorship characteristics in scientific publishing in the field of climate change adaptation and resilience research, to help provide insight on these issues.

The questions that guided this work are:

- What are the shape and composition of research networks in the field of climate change adaptation and resilience?
- Are southern actors and/or women actors becoming more central over time? Have some of these actors become key connectors?
- Are multi-disciplinary and multi-country collaborations impacting network topologies and especially positions for southern and/or women actors? (At this stage, the question of how these are collaborations impact network topologies could not be addressed.)

The rest of this introductory section will briefly outline methodological notes necessary to contextualise the findings presented in Section 2.

1.2 Methodology

1.2.1 Constitution of the CCARD publication set based on UN Sustainable Development Goals publication sets

A list of publications resulting form past funding was provided by IDRC was used as a starting point to delineate a global set of climate change adaptation and resilience-related publications from Scopus records. This list cut across multiple subfields from the Science-Metrix classification scheme, and only partially overlapped with the prior delineations of climate change adaptation research identified by Science-Metrix. Given the orientation of the IDRC and FCDO programs towards research with a strong development work component, the publications were examined for their distribution in the publication-level implementation of the United Nations (UN) Sustainable Development Goals (SDG) research

mappings previously defined by Science-Metrix for UNESCO and the European Commission.[2] The purpose of mapping research to SDGs was to be able to classify research according to its thematic alignment with the development issues. For example, a research article on the topic of coastal erosion in an African country is more likely to contribute to development efforts falling under SDG target 13.1 ("Strengthen resilience and adaptative capacity to climate related hazards and natural disasters in all countries") than a publication on the astrophysics of solar winds, and would therefore be classified as thematically aligned with this target and included in a publication set on this target. Nevertheless, Science-Metrix does not possess any data on whether publications have been used as input within actual development projects or initiatives – this classification simply captures thematic alignment with the SDGs. Using this approach, it was found that an appreciable share of IDRC-FCDO funded publications fell within a combination of four SDGs: Zero Hunger (2), Clean Water and Sanitation (6), Climate Action (13), and Life on Land (15).

To further narrow the pool of publications amongst which to position IDRC-FCDO publications, Scopus records were additionally filtered to include only those articles including a clear component of fieldwork in upper-middle income, lower-middle income, or low income countries (as defined by the OECD ODA list[3]; shortened to the "Global South" or "South" from here on). This requirement was operationalized by filtering out publications that did not include at least one author from a Global South country; or that did not mention a Global South location in their title, abstract or keywords. Identifying publications based on research conducted in a Global South location was performed using searches in the titles, abstracts and keywords of Scopus records based on a list of 12,000 global city, region and country names.

The result of this delineation work has been a publication set capturing a thematic focus on climate change adaptation and resilience (CCAR), which includes a methodological bias towards international development (adding a D to the abbreviation). The abbreviation CCARD will be used below to refer to this core publication set.

1.2.2 Alternative publication set: Topic Cluster 438

To determine whether specific and unique insights can be gained at the level of smaller research fronts (that cannot be gained by starting analyses at the level of the overall publication set), the analyses conducted on the CCARD publication set were also repeated on a Scopus-defined publication set called a Topic Cluster. There are close to 100,000 Scopus publication sets that are called Topics, established on the basis of citation networks and labelled using text mining, that share a common intellectual interest. These Topics are aggregated into broader, higher-level areas of research called Topic Clusters. Topic Cluster 438 combines publications from 121 Topics on themes such as climate change adaptation, resilience governance, and disaster management. Note that articles by North authors and without mention of a location in the Global South **are included** in this Topic Cluster's publication set and were included

[2] Rivest, M. et al. (2021). Improving the Scopus and Aurora queries to identify research that supports the United Nations Sustainable Development Goals (SDGs) 2021. *Mendeley Data*. Elsevier BV. doi:10.17632/9SXDYKM8S4.4.

[3] https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/DAC-List-ODA-Recipients-for-reporting-2021-flows.pdf

in analyses, with the explicit goal of offering a contrast with the CCARD publication set and providing some measure of the impact of this filter on the findings.

1.2.3 Production of networks of authors, based on co-authorship

Networks can be produced by examining two types of publication-level connections between researchers: citations and co-authorship links (a link is established when two researchers are authors together on one or more co-publication(s)). For this project, co-authorship networks were used. Co-authorship links represent strong collaborative ties, in contrast to citations which can be given by a certain author to other authors theoretically unknown to her. Citations, however, capture a broader definition of scientific influence and impact.

Authors in the network were characterized using the Scopus Author Identifier (AUID), a unique number that matches authorship to a group of documents. While some level of error in automated and semi-automated identifier curation has been shown in the past, the use of the identifier to characterize large aggregates (composed of 1,000+ publications) has been shown to be reliable.

It was possible to infer the gender of most researchers through information linked with AUIDs and the NamSor™ automated algorithm (see section 4.3.1.4 in the Appendix for details). Likewise, the country of affiliation for each AUID could be determined from the address information listed in publications authored. Note that gender could not be reliably inferred for more than 180,000 authors out of the 772,000 authors examined in this study, and inability to infer gender tended disproportionately affect Asian authors. Additionally, as some authors changed affiliation in time, North or South status was based on an AUID's affiliation(s) within a region with the greatest number of publications. North or South affiliation status could not be determined for more than 18,000 authors.

1.2.4 Centrality

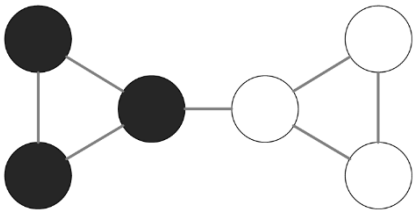
Betweenness centrality of specific groups of authors (women, men, affiliated to an institution located in the Global South; or located in the Global North) were calculated within the networks or subnetworks of interest. Within the specifications of this project, betweenness centrality measures the extent to which authors, or groups of authors, are sought out as collaboration partners. As of 2021, scientific research has become an intensively collective and collaborative practice, and it can be argued that scientists and scholars' productivity and status move up jointly with their capacity for collaboration and networking. Of course, there are also exceptions, and it is certainly possible that authors primarily working alone and publishing single-author articles make major contributions to any given field.

Centrality measures, including the author-level averages we use here, are crucially determined by the sheer volume of publications they contribute to a network. The more publications an author has, the greater likelihood they have (in principle) to have a long list of collaborators, and the greater likelihood they have to collaborate repeatedly with these authors. From prior experience, we expect publication volume and its immediate ramifications to be the most important explanatory factor behind high centrality scores by far.

To avoid extreme skewing effects in this measure which would have greatly concentrated centrality within a minor subset of authors, only authors with six publications or more in the CCARD publication set were retained for this particular analysis. In visual representations of the networks, a threshold of five publications or more was used instead, to maximise usage of available data.

1.2.5 Heterophily

Homophily is the tendency of individuals to associate and collaborate with similar others (Women with women; South author with South author). For example, if researchers collaborate more with other researchers from their country than with researchers from other countries, then the network is homophilic with regard to the country. Heterophily is the tendency of individuals to associate and collaborate with different others (women with men; North author with South author). Heterophily is thus a measure of integration. Traditionally, heterophily is calculated by comparing the average path length of actors that share a characteristic to the average path length of all actors in the network. In the following example, the average path length between any two actors is 1.8, but the average path length between white actors (or between black actors) is 1, meaning that the network is homophilic in regard to the colour of actors.



However, in the case of networks with disconnected components, this definition of heterophily is not useful because it becomes impossible to compute the average path length of disconnected actors. This was an issue for this project, because many of the networks involved high numbers of disconnected actors. Consequently, a modified version of this indicator had to be developed to provide insights on homophily in the networks. For a given network, the frequency of collaborations for each type of pairs (e.g. of themes, sectors, or countries) is computed based on the total number of collaborations in the network. Then, the number of collaborations of each pair in a model network of the same size (e.g. same number of actors and collaborations, but where co-publication links systematically follow the proportions of AUIDs in the women/men or southern/northern categories) is calculated to obtain a reference level.

It should be noted that analyses cannot determine whether 'integration' in co-publication opportunities is a positive or negative contributor to gender equality. For example, analyses cannot determine if women or South authors helping each other in relatively exclusive groups is a more efficient method to secure their advancement than full integration in the main network.

1.2.6 Authorships

In addition to authors and publications, an additional unit used in analyses is authorships. The concept of authorship captures single author-level contributions to publications. For example, a two-author article leads to two authorships. An author that has contributed to 10 publications in the CCARD set is

associated with 10 authorships. There are more authorships in the publication set than either publications or authors. Using this unit is useful to obtain paper-level averages of distributions among authors.

1.2.7 Additional analytical groups

While centrality and heterophily measures were produced for four main analytical groups and associated publications, these four groups were further intersected with a variety of other breakdowns to refine the analyses. Three dimensions are used as explanatory variables for author centrality and integration: degree of interdisciplinarity, multidisciplinarity and North-South co-publication in publication sets. These help address the main research questions of this study; namely, they address the impact of cross-disciplinary and North-South cooperative funding on women- and South-author positions and integration. Additional second-order dimensions were also implemented to qualify and nuance observations obtained. These include isolating publications by LMIC and LIC countries to help qualify the impact of North-South co-publications; isolating impacts related to open access (OA) publishing; and adding layers of interpretation by differentiating publications by the relative intensity of women participation in their authorship.

1.2.8 Main interpretation strategies

This report was written by focusing on differences between centrality and heterophily between groups from different subnetworks. In analyses that showed interesting differences (e.g. highly multidisciplinary publications, publications with a high share of women authorships), it may signal that the dimension captured by the breakdown drives differences in centrality or integration within the full network. To confirm this, more advanced analyses controlling for potential confounders would be needed but were beyond the scope of the current report.

Visualizations of the CCARD network were produced to represent the positions of authors in the network. This involved setting a threshold of five publications in the CCARD set by author, which filtered out between 91% and 95% of the (roughly) 772,000 authors initially present in the network (depending on the breakdown). Despite this drastic filtering effect, the approach still amounts to a relatively comprehensive take on network representation. Common filtering strategies in network analysis involve focusing on a top 50 or 500 authors or institutions in a network, an even more restrictive and incomplete approach. Authors with less than 5 publications can be considered to have made more transient contributions to the field, or even to science more broadly. At the same time, there is no doubt that established researchers benefit enormously from the small, discrete efforts of a great number of graduate students and other early career researchers that do not retain durable positions in science.

1.2.9 An initial discussion of limitations to keep in mind while reading

Important note for interpretation regarding targeting research conducted in Global South locations or by southern authors: IDRC-FCDO publications tend to include a component of fieldwork in Global South regions. This component was transformed into a requirement for northern authors when building the publication set used in this study. Publications with at least one southern author were automatically included in the CCARD set, for their part (provided they fell within the SDG dimensions of interest). While this helped increase the precision of the publication set, the upshot of

this design decision is that it may have favored the inclusion of publication of southern authors and the exclusion of publications by northern authors.

As explained in section 1.2.1 above, the publications (and by extension, associated researchers) included in this analysis were selected for a clear focus on fieldwork in the Global South through a keyword search or by filtering for the presence south-based authors. **This thematic operationalization might have introduced a bias in favor of southern authors, at least in terms of their relative representation within the network**. The potential unintended ramifications of this methodological choice should be remembered in interpreting and reproducing elsewhere the findings on South vs North author centrality presented below.

It must also be kept in mind that it was not possible to obtain definitive evidence that multidisciplinarity and North-South collaboration are causal drivers of changes in network topology (especially with regard to the centrality of south and women actors in the CCARD publication set or in the Topic Cluster 438 publication set). Obtaining more definitive evidence of this would require the use of advanced statistical modeling, which was out of the scope of this study. See section 3.3 for details.

2 Results

2.1 Are women actors more central over time? Have women actors become key connectors?

2.1.1 Findings on women centrality and integration in the overall dataset

In keeping with the broad analytical lines presented above, the betweenness centrality of women authors was calculated in a number of dimensions, and compared against the centrality of men authors. Additionally, shares of those women and men authors whose centrality was so high, it placed them within the top decile of most central actors in a given network or sub-network, were also computed. These percentage shares can capture the proportion of women or men authors that are "key connectors".

Heterophily computations provide a sense of the extent to which co-publication opportunities across various groups (e.g., based on gender or geographic location) are equally distributed among men and women; whether women are included or excluded from specific groups (e.g., north or south actors) in performing collaborative research projects (which, by and large, make up a major mode of doing research at this point in time).

To better tease out any potential differences in centrality and heterophily between women and men authors, breakdowns of publications have been assembled. Publications were separated in bins according to the proportion of their (co-)authors that were recorded as women. Publications with a high proportion of women authorship were those where this figure was recorded at 75% of more, while those with a low proportion had shares of 25% or less of authorships held by women. Of course, in the specific case of the bin of publications with a high proportion of women authorships, women homophily is expected to be dominant, and therefore the assessment of heterophily must be interpreted from a slightly different perspective as in the rest of this report.

To start off (Table I), it can be noted that about one third (33%) of authorships in the CCARD thematic set were held by women for the most recent period considered (2015–2019). This share of authorships was larger than the share of women found amongst unique (i.e., distinct) individual authors in the field (26%; against 50% of authors identified as men and the remaining 24% of authors having been classified as of unknown gender). The share of total authorships held by women has slowly increased in time, and was previously at 31% in 2010–2014 and 29% in 2005–2009.

Findings on betweenness centrality show that women authors lag behind men authors in the full CCARD set, on average. Centrality scores for women were 29% below those of men in the recent period. A share of 8.0% of women were amongst the top decile of most central authors in the field, against 11.2% for men.

These recent achievements of women authors on centrality nevertheless do amount to improvements in relative positions to men, in comparison to the gaps recorded in prior periods. Women's gap on average centrality was -38% in the two prior periods (2005-2009 and 2010-2014); shares of highly central women were of 7.3% in the first period (against 11.2% again for men).

Table I Centrality and integration of women in the CCARD publication set, 2005–2019

	Thematic set - overall			High women particip.			Mid women particip.			Low women particip.		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of publications	65,007	117,378	202,883	5,051	8,835	14,730	25,689	52,766	101,185	32,524	53,526	83,612
Share of women amongst authors	29%	31%	33%	93%	92%	91%	44%	44%	44%	7%	8%	9%
Growth index publications with 2+ women authors		1.19	1.13		1.00	1.01		1.12	1.08		1.53	1.27
Network centrality and positions												
Relative difference in centrality degree, women authors against men authors	-37.9%	-37.8%	-28.9%	*6.7%*	*372.5%*	*139.0%*	2.1%	-2.8%	1.9%	-63.0%	-67.9%	-65.8%
Share of highly central women authors	7.3%	8.0%	8.0%	*11.1%*	*10.9%*	*11.5%*	11.1%	10.6%	10.9%	4.1%	4.1%	3.6%
Share of highly central men authors	11.2%	11.0%	11.2%	*3.1%*	*6.6%*	*5.8%*	9.8%	10.1%	10.0%	10.9%	11.0%	11.2%
Women-Men integration in co-authorship												
Heterophily Women authors-men authors	0.90	0.92	0.91	1.09	1.09	1.09	1.09	1.03	0.97	1.01	1.03	1.02
Homophily women authors - women authors	1.26	1.20	1.18	0.99	0.99	0.99	0.86	0.94	1.05	0.89	0.76	0.81
Homophily men authors - men authors	1.04	1.03	1.04	0.24	0.20	0.25	0.94	0.98	1.02	1.00	1.00	1.00

Note: High women participation publications: 75% of more of women authorship. Mid women participation publications: 75 to 25% of women authorship.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

Heterophily numbers in the full CCARD set show that there is a certain tendency for women to co-author publications preferentially with other women (propensity of heterophilic pairs of 0.91; while for women-women homophilic pairs the score is 1.18).

Key finding 1: Over the analysis period, women have increased their relative centrality (while remaining at a gap to men) and slightly increased their participation in authorship. Women's average centrality relative to men has moved from -38% to -29% between 2010-2014 and 2015-2019. They account for 33% of publication authorships (defined as author-publication pair, with a single author or publication being associated with multiple authorships) in the most recent period, against 31% and 29% in the two prior periods.

Authors have a slight tendency to collaborate with authors of the same gender, and this has been stable over time (heterophily measured at around 0.90).

2.1.2 Women centrality in publications with high and low intensity of women authorship

The graphical mapping of the network (Figure 1) supports the observation that men and women authors do act together within a single network (heterophily is below expected by a small margin) and that despite a gap in average centrality for women, women are positioned in the core of the network and not just its periphery. This network mapping is based on single authors with 5 or more publications in the CCARD field. While co-publication links between authors are not represented to ease interpretation, if heterophily had been low, men and women authors would have been segregated in distinct clusters of the network graph. If centrality would have been much more unevenly skewed (towards men authors for instance), the central region of the graph would have been heavily or exclusively populated by these authors. The graphs produced for the three periods of interest are all highly similar on these dimensions, indicating that these basic coordinates have not drastically changed over time (illustration not shown) even though progress was noticed based on network indicators (Table I).

Turning to the analysis of breakdowns by levels of women authorship in publications, the set of publications with the highest level of participation of women recorded roughly 92% of authorships to be held by women. Within this set of publications, women authors expectedly have high average centrality scores, collectively. The average score of women was 140% above that of men in the recent period, and much above that in the period just prior. It should be noted here, however, that this comparison is based on extremely low centrality numbers, and therefore should be interpreted with caution. Indeed, the main conclusion here may be that the subset of publications with high women authorship resulted in a particularly flat network, with very little in the way of hierarchies in centrality otherwise observed in the networks or sub-networks produced in the CCARD field. Complicating matters yet further, shares of highly central authors by gender group showed men to perform comparatively well in this bin of publications, despite amounting to only 5% of authors in this network. A share of 6% of men were amongst the top decile of authors on centrality, against 12% of women. In other words, research with a strong women presence still appeared to involve core men actors nonetheless.

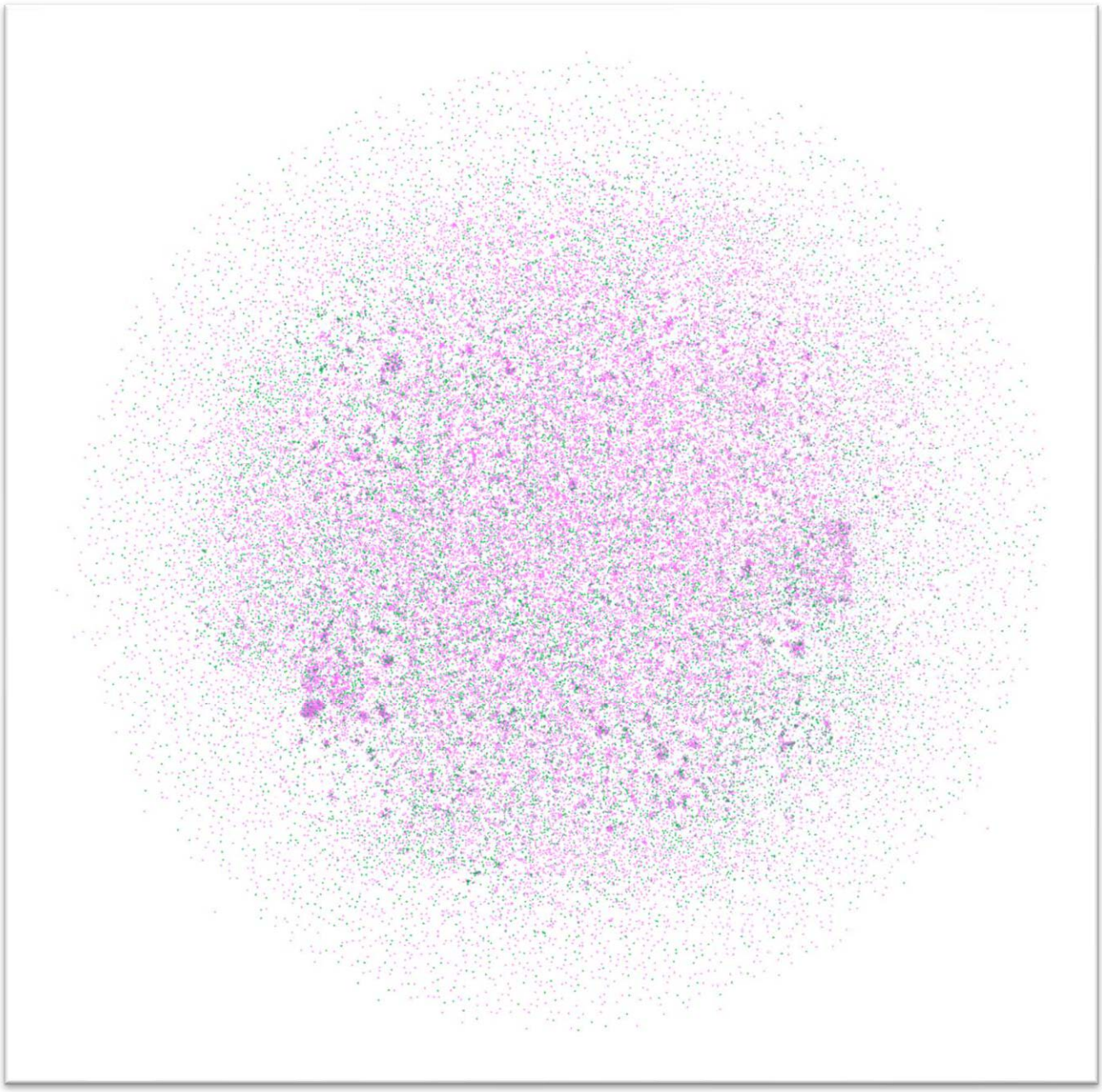


Figure 1 Network mapping of women and men authors in the CCARD field, 2015–2019

Note: Women colored in green, men in purple. Placement of author on the mapping is indicative of centrality. Positions of authors relative to one another is indicative of heterophily/homophily.

Source: Prepared by Science-Metrix using Scopus (Elsevier) data and Gephi for visualization

In the breakdown of publications with middle levels of participation by women (44% of authorships here are held by women on average), the average of centrality scores for women is roughly the same as that of men. This bin contributed to almost 50% of publications in the CCARD field. This observation, combined with the prior observations on centrality in the bin of publications with high women participation, indicate that the gaps in centrality of women to men in the field can be fully tracked back

to those papers in the bin with low intensity of women authorship. Here, only 8% of authors are women, their average centrality is 66% below that of men; and only 4% of women fall within the top decile of most central authors. This shows that as women keep closing the gap to men, women should naturally become as prominent as men in CCARD research.

Key finding 2: Men tended to have high levels of average centrality even in those publications with high levels of authorship by women. Even in the small subset of publications (less than 10% of the field) where women provided 75% or more or authorships, contributing men tended to make up a comparatively high proportion of the network's most central authors (6 or 7% in the two recent periods, against 11-12% for women).

2.2 Are southern actors more central over time? Have southern actors become key connectors?

2.2.1 Findings on south author centrality and integration in the overall dataset

Looking at Table II, South authors appear to be behind North authors for their relative centrality in the network of CCARD research. **It bears repeating that this is this despite a dataset that has a potential bias in favor of South actors per its construction**. In the 2015–2019 period, the average betweenness centrality measurements of South authors were 33 % below those of North authors. Amongst North authors, 12.4% recorded betweenness centrality scores that put them amongst the decile of overall authors with the highest centrality scores. This score was 8.9% for South authors. Leads for North authors over South authors on these two dimensions had moderately increased over time, moving from to the 2005-2009 period to the 2010–2014 and 2015–2019 periods. For instance, the gap of South authors on average betweenness centrality has moved from -20% in the initial period to the already mentioned score of -33% in the most recent period.

In the CCARD thematic publication set, the network was found to tend more towards heterophily than homophily. North–South co-publication links were observed 8% more than expected by chance. However, like the centrality of South actors, the prevalence of North–South heterophilic pairs has been decreasing in time, as it was measured at 1.23 for 2005–2009, 1.14 for 2010–2014, and 1.08 for 2015–2019. South authors, in particular, increased their propension to co-author publications with other South authors over time, as the prevalence of South-South homophilic pairs going up from 0.67 to 0.88 from the first to last period.

The graphic representation of the CCARD network for the 2015–2019 period, with authors colored by South- and North- affiliation, is shown in Figure 2. This graph provides a somewhat clearer representation of the centrality gap of South authors (than the gender-based graph presented previously), and the difference in centrality is even clearer in the graph from the 2010–2014 period (Figure 3). Despite a higher heterophily score for co-publications between southern authors with northern authors than women authors with men authors, authors in Figure 2 appear more clustered than in Figure 1. The likely reason for this is that the network of southern and northern authors is shaped by country-level dynamics as a confounding factor, with countries likely acting as attractors for clusters of either southern or northern local authors.

Table II Centrality and integration of South authors in the CCARD publication set, 2005–2019

	Thematic set - overall			North-south co-pubs			LMIC publications			LIC publications		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of publications	65,007	117,378	202,883	17,472	32,495	64,040	22,204	40,297	74,735	2,397	4,321	8,060
Network centrality and positions												
Relative difference in centrality degree, South authors against North authors	-20.1%	-23.6%	-33.2%	-34.3%	-40.1%	-44.3%	19.8%	10.9%	3.4%	57.8%	70.3%	29.3%
Share of highly central North authors	11.1%	11.6%	12.4%	12.0%	12.8%	13.4%	9.5%	10.0%	9.5%	7.7%	8.3%	8.6%
Share of highly central South authors	9.5%	9.2%	8.9%	8.2%	7.6%	7.6%	10.3%	10.2%	10.6%	12.9%	12.4%	12.1%
North-South integration in co-authorship												
Heterophily North authors - South authors	1.23	1.14	1.08	1.39	1.26	1.18	1.23	1.14	1.07	1.16	1.07	1.02
Homophily North authors - North authors	0.84	0.91	0.95	0.71	0.82	0.89	0.78	0.88	0.94	0.79	0.91	0.98
Homophily South authors - South authors	0.67	0.79	0.88	0.48	0.61	0.73	0.76	0.85	0.92	0.88	0.94	0.98

Note: LMIC: low -middle income countries. LIC: low income countries.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

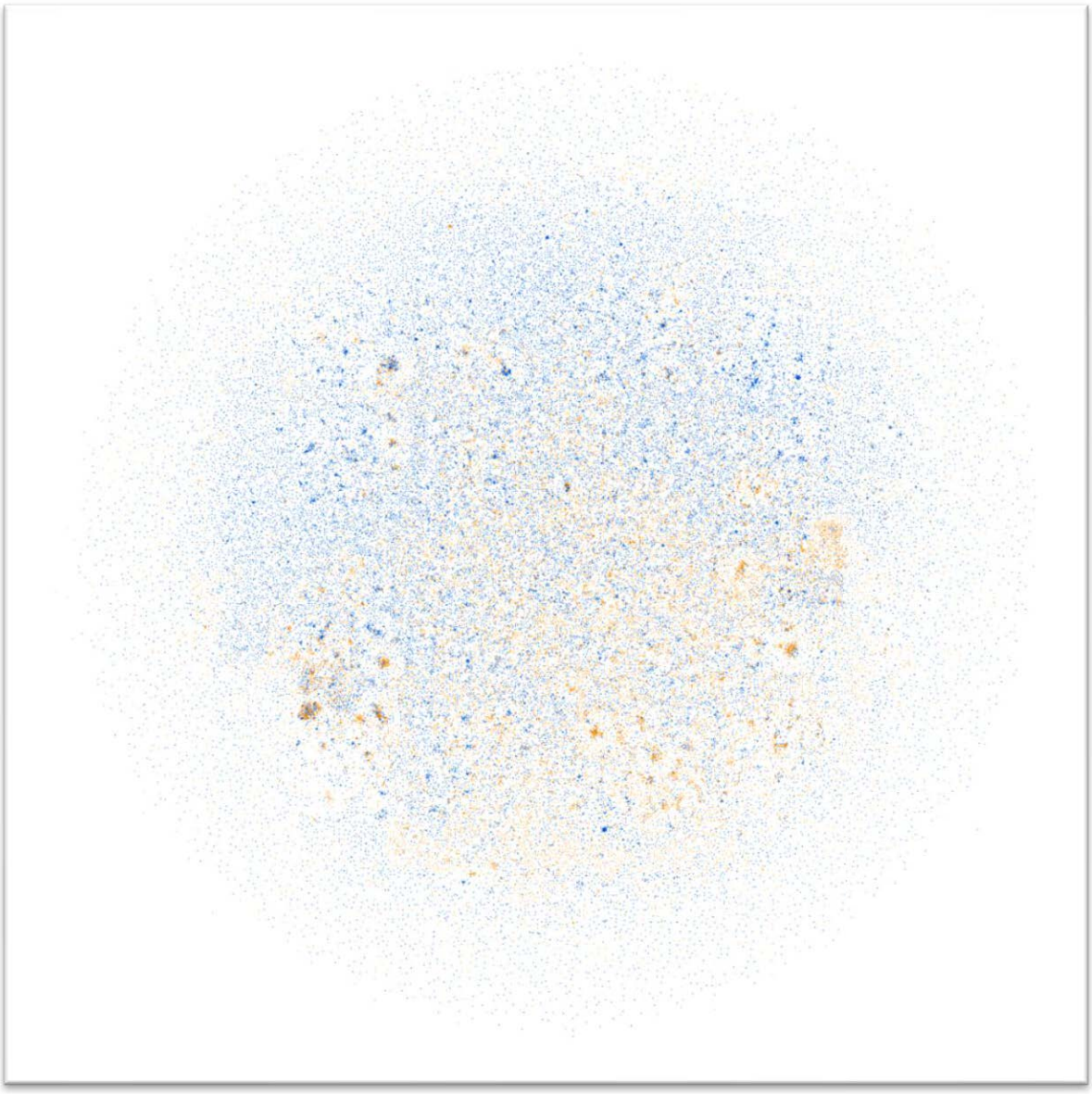


Figure 2 Network mapping of South and North authors in the CCARD field, 2015–2019

Note: South authors are in blue, North authors are in orange. Placement of author on the mapping is indicative of centrality. Positions of authors relative to one another is indicative of heterophily/homophily.

Source: Prepared by Science-Metrix using Scopus (Elsevier) data and with visualization by Gephi

What these findings show so far is that South authors were well present and integrated in the field. Numerically, they make up a smaller portion of researchers in the field (44%) than North authors (51% - data not shown), despite a thematic requirement to conduct fieldwork in Global South countries, or be affiliated with a Global South institution, in the definition of the data set used. This situation of almost numerical equality evident in Figure 2. There is also higher than expected heterophily in terms of co-authorship of publications between North and South authors, indicating that South authors are not excluded from research teams and partnerships (although this does not tell us whether they are integrated as equals or in more subordinated roles). What is also clear, however, is that North authors tendentially

(but moderately so) hold more central positions in the network than South authors. This trend has been reinforced in time. Heterophily has decreased in time, with South-South homophily in co-authorship particularly making gain through that development.

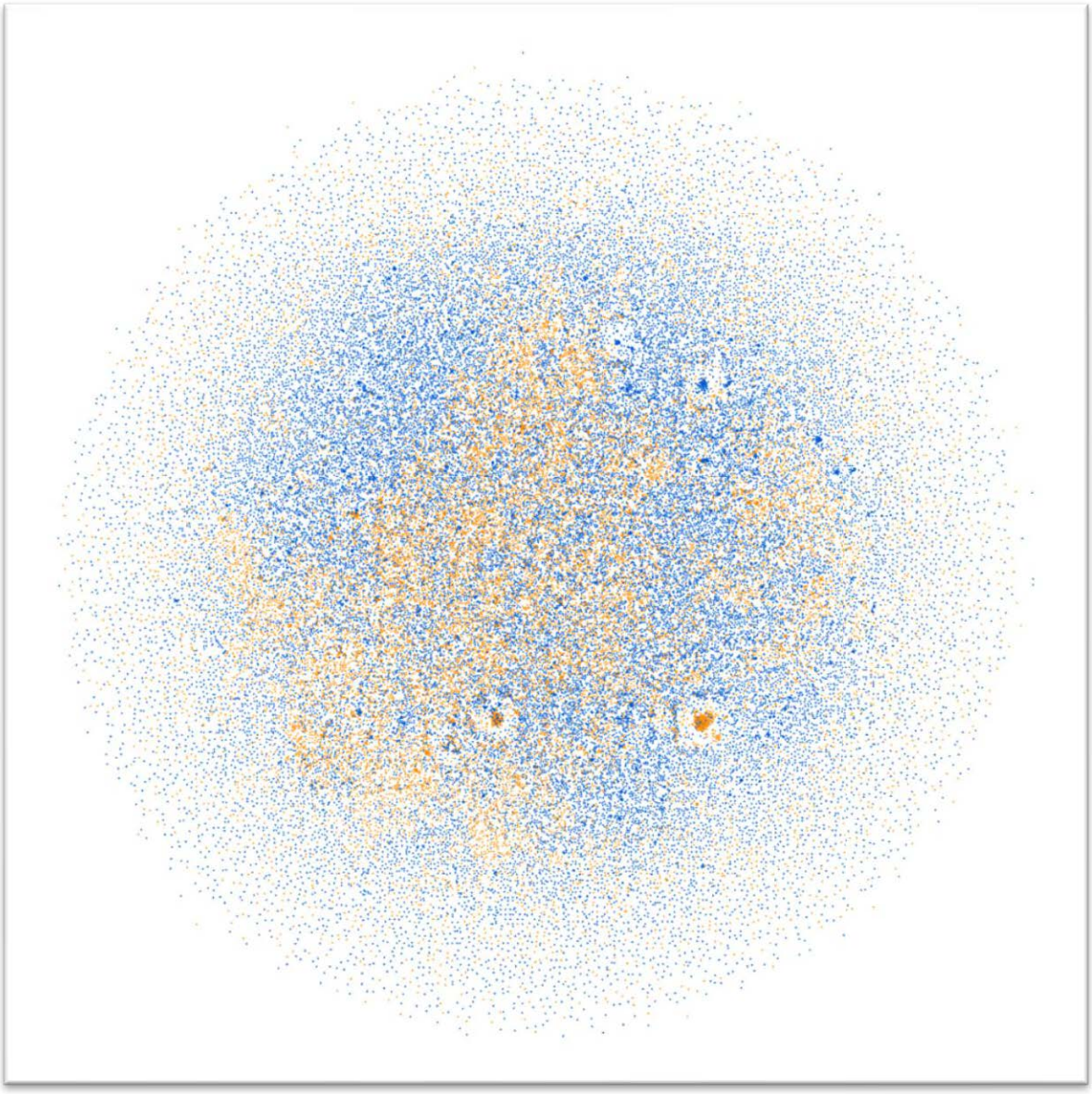


Figure 3 Network mapping of South and North authors in the CCARD field, 2010–2014

Note: South authors are in blue, North authors are in orange. Placement of author on the mapping is indicative of centrality. Positions of authors relative to one another is indicative of heterophily/homophily.

Source: Prepared by Science-Metrix using Scopus (Elsevier) data and with visualization by Gephi

Key finding 3: South authors have a lower average centrality compared to North authors, and the gap has increased over time. Despite the design decision to focus on publications with clear regional engagement in Global South countries, South authors appear to be behind North authors in terms of their relative centrality in the field's network. The gap in centrality has actually increased over time (from -20% to -24%, and then to -33%).

While the network saw heterophily on the North-South dimension, there has been a tendency towards more homophily in time (heterophily moving from 1.23 to 1.14 and then 1.08). That is, North-South co-authorships tend to represent a diminishing share of the overall amount of authorship relationships in the field.

2.2.2 South author centrality in North-South collaborations and publications by LMIC/LIC researchers

Comparing the findings just presented above against equivalent indicator measurements in relevant analytical breakdowns can add an additional layer of interpretation. Findings on South author centrality in North-South co-publications can inform on how these international collaborations contribute to the trends identified above. Findings for subsets of publications from LMIC and LIC countries (albeit potentially written as co-publications with authors from other countries) may reveal different patterns, given how scientific production in the South is expected to be crucially shaped by the larger national publication outputs of upper middle-income countries. These additional findings are presented in the right-side columns in Table II.

Considering specifically North-South co-publications in the thematic set, it appeared this group of collaborative papers might be linked to the inequalities in centrality between North and South authors in the full network. Gaps in average centrality for South authors against North authors were more pronounced in this subset (with a difference measured at -44% in the recent period, against -33% in the full thematic set). Even more so than in the full thematic set, Southern authors made up a smaller proportion of the top decile of highly central authors in the field (7.6% in the last period, against 8.9% in the full thematic set). Gaps for South authors against North authors all increased in time. Unsurprisingly, and most likely by definition, heterophily was higher in this subset of publications (1.18 against 1.08 in the recent period).

Interestingly, however, the breakdowns for LMIC and LIC publications saw higher proportions of North-South co-publications (39% and 68%, respectively) than in the full CCARD set; while simultaneously recording average centrality measurements that were favorable to South authors. Most strikingly, in the comparative small set of LIC publications (8,000 units in the recent period; against 200,000 in the full thematic set), South authors' average centrality was 30% above that of North authors. A share of 12% of South authors were amongst the top centrality decile in the field, against 9% for North authors. LMIC publications also saw a lead in centrality for South authors, although not as pronounced as in the LIC publications. That said, in the LIC publication sets, South authors leads were decreasing in time (for instance, from 58% to 29% in average centrality between the first and third periods). The smaller leads in the LMIC set were roughly stable in time considering the shares of author in the top centrality decile, but decreasing for the average centrality indicator (from 20% to 3%).

Although the comparison between North-South publications on the one hand and LMIC/LIC publications on the other is imperfect, this exercise nevertheless provided signal that North-South co-publications may have quite different outcomes on South author centrality in the CCARD network depending on the exact country involved in the collaboration. It appears North-South co-publications may have more positive effects on South author centrality if these authors are from LMIC or, even more so, LIC countries; but presumably much less so in UMIC countries, although this was not directly tested.

Key finding 4: North-South co-publications have been associated with a decrease rather than an increase in South author centrality, and the resulting gap has increased over time (from -34% to -40% and then -44%; note that this trend mirrors a similar trend in the main dataset)**.** Considering only publications with at least one contribution by a low income country-based researcher, South authors leads were decreasing in time (difference in average centrality for South authors went up from +58% to +70% between 2005–2009 and 2010–2014, but went down to do +29% in 2015–2019). In lower-middle income country publications, the difference in average centrality for South authors against North authors decreased from +20% to +11%, and then almost reached parity at +3% in the latest period.

2.3 Have multi-disciplinary collaborations impacted network positions for women actors?

The percentage shares of authorships that were identified as held by women in the overall thematic publication set went from 29 % in the 2005–2009 period to 33% in the most recent period (2015–2019 – see Table III). By and large, women's shares of authorships were comparable in the subsets of highly interdisciplinary and highly multidisciplinary papers. It can be noted, however, that publications that were both interdisciplinary **or** multidisciplinary **and** included at least 2 women as co-authors grew in volume more slowly (by a slight gap) than other publications with at least 2 women as co-authors (growth indices of 1.07 or 1.09 in the recent period, compared to 1.13 in the full CCARD dataset).

Interdisciplinary and multidisciplinary publications may be conducive to increased network centrality for women authors. While in all period-breakdown combinations women were at a gap to men in terms of centrality, in most cases this gap was smaller than the one measured in the full thematic set. For instance, considering the recent period (2015–2019) and the set of publications falling amongst the top decile of most multidisciplinary papers in their subfield ($HMP_{10\%}$), the women to men gap in centrality was -17%, against an overall thematic figure of -29%. The gap was somewhat larger in the network made of publications from the top decile in interdisciplinarity ($HIP_{10\%}$), with a difference of -21% in centrality scores. Measurements in terms of the share of women and men authors falling in the top decile of authors with the highest centrality scores broadly followed the same pattern, except for highly interdisciplinary papers from the two earlier periods, where the share of highly central women authors came close to equaling that of men (10% to 11%).

Changes in centrality over the periods of interest may warrant the conclusion that the conduct of multidisciplinarity, and to a lesser degree, interdisciplinary projects may support gains in centrality for women authors. While the overall CCARD field has seen increasing relative centrality for women authors moving to the recent period (from -38% in the two earlier ones to -29%), a bigger gain has been made

specifically in the network of highly (10%) multidisciplinary papers. This gain was recorded as moving from -33% to -17%. While other gains in time have been more modest (keeping roughly with the full field's effect size), these gains nevertheless maintain the already comparatively more favorable centrality of women authors in interdisciplinary and multidisciplinary publications. Since the group of authors (both women and men) appearing on highly multi- or inter-disciplinary papers may not be representative of the populations of researchers in the overall CCARD network, we cannot confirm whether the above observations are due to multi- and/or inter-disciplinary research increasing the centrality of women relative to men, or if they are due to more central women more frequently engaging in this type of research.

Multidisciplinary and interdisciplinary papers did not appear to alter the pattern of co-publication between men and women, as shown by heterophily findings that were roughly comparable to those of the overall thematic set.

Table IV takes a different angle on the correlation between women authorship and cross-disciplinary publication. The $HMP_{10\%}$ and $HIP_{10\%}$ shares of publications amongst publication in the three bins of publications differentiated by level of women participation were recorded. Publications with a high proportion in women authorship were particularly likely to be highly multidisciplinary (18%), a small lead over the figure in the overall set (16%). The trade-off, however, was that the same bin of publications also saw a much lower share of highly interdisciplinary publications (11%, instead of 16% in the overall set). This observation provides further support that women authors in the CCARD field may gravitate towards collaborative projects involving partners from a greater diversity of disciplines than what is common practice.

Key finding 5: The position of women relative to men has been improving over time across all sets of cross-disciplinary publications considered, although the difference in average centrality remains negative. For instance, in highly multidisciplinary publications, those negative differences have reduced from –33% (two earlier periods) to -17%. Women held roughly as many authorships in the subsets of highly multidisciplinary and interdisciplinary publications as in the field overall.

Table III Centrality and integration of women in CCARD cross-disciplinary publications, 2005–2019

	Thematic set - overall			HMP (10%)			HMP (30%)			HIP (10%)			HIP (30%)		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics															
Number of publications	65,007	117,378	202,883	5,340	12,134	25,945	15,296	32,735	65,721	5,912	14,630	30,782	15,104	36,376	74,754
Share of women amongst authors	29%	31%	33%	28%	31%	32%	29%	31%	32%	31%	33%	35%	31%	33%	34%
Growth index publications with 2+ women authors		1.19	1.13		1.16	1.07		1.16	1.09		1.15	1.07		1.15	1.07
Network centrality and positions															
Relative difference in centrality degree, women authors against men authors	-37.9%	-37.7%	-28.9%	-32.0%	-32.6%	-16.8%	-35.8%	-33.5%	-26.7%	-26.8%	-24.3%	-21.1%	-43.5%	-29.2%	-27.9%
Share of highly central women authors	7.3%	8.0%	8.0%	7.7%	8.6%	8.9%	7.5%	7.9%	8.1%	9.9%	9.9%	8.7%	8.6%	8.5%	8.4%
Share of highly central men authors	11.2%	11.0%	11.2%	10.8%	10.7%	10.7%	11.0%	11.1%	11.1%	10.7%	10.7%	10.8%	10.9%	10.9%	10.9%
Women-Men integration in co-authorship															
Heterophily Women authors-men authors	0.90	0.92	0.91	0.93	0.92	0.95	0.92	0.93	0.94	0.91	0.91	0.90	0.91	0.92	0.90
Homophily women authors - women authors	1.26	1.20	1.18	1.18	1.17	1.13	1.20	1.17	1.13	1.21	1.17	1.15	1.23	1.17	1.17
Homophily men authors - men authors	1.04	1.03	1.04	1.03	1.04	1.02	1.03	1.03	1.02	1.04	1.04	1.07	1.04	1.04	1.05

Note: HMP: Share of highly multidisciplinary (top decile or top three deciles) publications; HIP: Share of highly (top decile or top three deciles) interdisciplinary publications.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

Table IV Cross-disciplinary publications by intensity of women authorship in CCARD publications, 2005–2019

	Thematic set - overall			High women authorship			Mid women authorship			Low women authorship		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of papers	65,007	117,378	202,883	5,050	8,835	14,730	25,686	52,763	101,168	32,523	53,517	83,603
Achievements in cross-disciplinarity research												
Growth index HMP		1.15	1.17		1.28	1.17		1.14	1.16		1.14	1.17
HIP	11.8%	13.5%	15.8%	7.4%	9.5%	11.1%	12.6%	14.3%	16.6%	11.5%	13.2%	15.5%
HMP	12.8%	14.2%	15.6%	13.9%	16.3%	18.3%	13.2%	14.6%	16.2%	12.3%	13.4%	14.4%

Note: HMP: Share of highly multidisciplinary (top decile) publications; HIP: Share of highly (top decile) interdisciplinary publications. High women participation publications: 75% of more of women authorship. Mid women participation publications: 75 to 25% of women authorship.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

2.4 Have multi-disciplinary collaborations impacted network positions for southern actors?

A presentation of descriptive findings can provide a first few general interpretative signposts before directly tackling the question of whether multi-disciplinary collaborations have impacted network positions for South authors. Highly multi-disciplinary publications (those falling either in the top decile or top three deciles in the distribution of multidisciplinarity scores) contained a higher share of international co-publications than the full dataset (for instance 53% for $HMP_{10\%}$ papers against 38% for the most recent period – see Table V). The difference on this dimension was roughly negated for inter-disciplinary papers. Much the same conclusions held for North-South co-publications. For instance, 45% of highly (10%) multi-disciplinary co-publications are also North-South co-publications, against 32% in the full dataset, or 34% for highly interdisciplinary papers – all of this in the most recent period.

Turning to centrality scores, multidisciplinary and interdisciplinary publications appeared to have been *negatively correlated with centrality* outcomes for southern actors' positions in the two first periods considered. For example, we observe a greater centrality gap between north and south actors in the 10% HIP relative to all papers, especially in the first two periods (Table V). Also, the share of North authors amongst the top decile of most central authors within the 10% highly interdisciplinary publications was 15%, compared to 6% for South author, between 2005 and 2009. These numbers compared to 11% and 10%, respectively, in the full data set.

That high interdisciplinarity is negatively associated with centrality of south actors at the aggregate level, what could be due to interdisciplinarity itself or other covariates. For example, South actors producing HIP might not be representative of all actors from the South in the full network. They might collectively have different characteristics. For example, women and less senior researchers have been shown to have a greater propension towards interdisciplinarity[4] and may thus be overrepresented in the 10% HIP. Since these groups are typically associated with smaller outputs per individual, they would naturally be less central in collaboration network; centrality correlates strongly with number of publications per individual. While the above would also be true for north actors, women and early career researchers from the South might have a stronger gap in production volume to their peers in the north than the North-South productivity gaps observed for male and senior researchers. If the above is true, the remaining question would then be whether it is the higher propension towards interdisciplinarity, or the lower productivity, of women and/or early career researchers that led to the observed differences in the North-South centrality gap for the 10% HIP. Other confounders could also be at play here. Identifying some of the "true" causal factors and their effect sizes would require more advanced statistical modelling (see section 3.3 and after).

Much the same observations as made above can be applied to multidisciplinary publications, down to the periodic change from a high gap of South authors on centrality moving towards a gap that is in line with what is observed in the full CCARD dataset in the recent period.

[4] Durning, Pinheiro and Campbell. 2021. Do Women Undertake Interdisciplinary Research More Than Men, and Do Self-citations Alter These Perceptions? Submitted to *Quantitative Science Studies*.

Table V Centrality and integration of South authors in CCARD cross-disciplinary publications, 2005–2019

	Thematic set - overall			HMP (10%)			HMP (30%)			HIP (10%)			HIP (30%)		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics															
Number of papers	65,007	117,378	202,883	5,340	12,134	25,945	15,296	32,735	65,721	5,912	14,630	30,782	15,104	36,376	74,754
Collaboration practices															
Share international co-publications	33.1%	34.2%	38.2%	50.4%	50.5%	53.0%	47.1%	47.7%	50.3%	37.6%	38.5%	40.9%	38.7%	38.4%	41.1%
Share North-South co-publications	27.2%	27.8%	31.6%	42.6%	42.1%	44.8%	39.6%	39.4%	42.4%	30.9%	31.0%	33.6%	31.9%	31.3%	33.8%
Growth index North-South co-publications		1.03	1.14		0.99	1.07		1.00	1.08		1.00	1.09		0.98	1.08
Share South-South co-publications	2.1%	2.6%	3.3%	2.4%	3.6%	4.7%	2.3%	3.3%	4.3%	1.6%	2.5%	3.7%	1.7%	2.4%	3.5%
Share North-North co-publications	3.7%	3.8%	3.3%	5.4%	4.8%	3.5%	5.2%	5.0%	3.7%	5.0%	5.1%	3.6%	5.1%	4.7%	3.7%
Network centrality and positions															
Relative difference in centrality degree, South authors against North authors	-20.1%	-23.6%	-33.2%	-69.6%	-24.7%	-32.7%	-36.1%	-20.4%	-29.6%	-84.3%	-54.4%	-36.9%	-63.8%	-26.6%	-35.0%
Share of highly central North authors	11.1%	11.6%	12.4%	15.0%	11.7%	11.7%	12.3%	11.5%	12.0%	15.0%	15.2%	12.2%	14.7%	12.5%	12.2%
Share of highly central South authors	9.5%	9.2%	8.9%	6.6%	9.0%	9.2%	8.6%	9.2%	9.0%	6.3%	6.6%	8.8%	6.7%	8.5%	8.9%
North-South integration in co-authorship															
Heterophily North-South	1.23	1.14	1.08	1.09	1.07	1.04	1.15	1.08	1.05	1.19	1.16	1.11	1.21	1.14	1.08
Homophily North-North	0.84	0.91	0.95	0.94	0.96	0.97	0.90	0.95	0.97	0.88	0.90	0.93	0.86	0.91	0.95
Homophily South-South	0.67	0.79	0.88	0.86	0.88	0.95	0.77	0.88	0.93	0.72	0.76	0.84	0.68	0.76	0.88

Note: HMP: Share of highly multidisciplinary (top decile or top three deciles) publications; HIP: Share of highly (top decile or top three deciles) interdisciplinary publications.

Source: Prepared by Science-Metrix using Scopus (Elsevier) data

It can be noted that in terms of integration in co-authorship, heterophily level was above expected in all cases, roughly on the same level as the heterophily levels measured in the overall publication set. This observation reinforces the view that opportunities for North-South collaborations can co-exist with a much more uneven distribution in centrality along the North-South divide. Recall, however, that the thematic dataset is biased, by construction, in favor of the global South.

Key finding 6: Although multiple signals do not fully converge in one simple observation, on balance, cross-disciplinary research likely had a negligeable impact on South author positions and potentially a negative impact on one dimension. Highly multidisciplinary publications tended to include more North-South co-publications (recording a share of 45% such collaborations, against 32% in the overall publications). Yet, over 2010–2019, South authors' centrality within highly cross-disciplinary publications followed closely the levels and trends found in the overall publication set. For example, the difference of South author centrality against North author centrality increased from -25% to -33% over those periods. The exception was for the most highly interdisciplinary research $HIP_{10\%}$, where South authors recorded a much larger gap in average centrality than North authors compared with the corresponding gaps recorded in the overall publication set.

2.5 Have North-South co-publications impacted network positions for women actors?

To assess the extent to which multi-country collaborations have impacted network positions for women actors, women/men betweenness centrality and heterophily were measured in the subset of CCARD publications that are North-South co-publications; that is, international co-publications where at least one author is affiliated with an institution from a "North country" (high income country in the OECD ODA list) and one is from a "South country' (upper middle, lower middle and low income countries, but excluding China, given the achievements of its science system). Findings were also produced for publications by aggregates of LMIC and LIC countries, to provide further details into the potential influence of specific levels of national scientific capacity on these findings. These LMIC and LIC publications were not required to be North-South co-publications, although North-South co-publications did account for 39% and 68% of these country aggregates' papers.

First, it is possible to notice that there are slightly less women amongst authors on North-South co-publications than in the overall thematic set (Table VI). The differences are small at two or three percentage points less (for instance 31% to 33% in the recent period), but they are constant across the three periods. Similar gaps in women participation are also found in LMIC and LIC publications.

The volume of publications with two or more women authors has grown at a faster pace in the subsets of North-South co-publications, LMIC publications and LIC publications; although the difference was more pronounced between the 2005–2009 and 2010–2014 periods than between the two more recent periods. For instance, the growth index was measured at 1.28 for North-South co-publications between the two earlier periods, against 1.19 in the overall set.

Roughly and at a high level, women authors' centrality was at comparable levels in North-South co-publications as it was in the overall set. Women's average gaps to men in centrality measurements were

very close in both groups of publications, different in the two earlier periods by only one or two percentage points. In the recent period, relative difference in centrality was measured at -29% in both publication sets. Considering shares of authors among women and men falling within the top decile of centrality scores in the CCARD field, women had a slightly decreased gap in the recent period (8.4% of women in North-South co-publications against 8.0% in the overall set; both comparing to about 11% of men). However, women's gap on this dimension was larger in the period immediately prior.

Heterophily measures in the North-South co-publication set show that the level of gender integration here is very slightly higher (0.93 in the recent period) than in the overall set (0.91; with this level of difference reproduced in prior periods). North-South co-publications reproduce a certain tendency of women to preferentially collaborate with other women.

From these observations so far, it can be concluded that North-South co-publications have likely not altered women's network positions in the CCARD field.

Findings from the LMIC and LIC aggregates are particularly interesting, especially for the latter group of publications. Women authors were collectively at a smaller gap to men here, for instance with a difference in average betweenness centrality reduced to -14% in the 2010–2014 period. This, despite the very high degree of North-South co-publications as a proportion of overall publications originating from this group of country (68% as mentioned above; with 80% of publications being international co-publications). These observations point towards the possibility that North-South co-publications might have differentiated impacts in different country groups, and follow-up work might compare outcomes for women in a more complex typology of collaborations (to take just one example, women in LICs may benefit differently from North-South co-publications than those in UMICs).

Finally, Table VII presents collaboration patterns in the bins of publications categorized by relative intensity of women participation in authorship. Highly interesting findings were found for the small bin of publications with the highest intensity of women participation. In these papers, on average, 91% of authors are women, much lower values were recorded for all forms of international co-publications. The share of overall international co-publications was only 18% (against 38% in the overall thematic set). Needless, to say, figures for North-South, South-South and North-North co-publications were accordingly lower than in the overall set as well. This observation might provide a signal (by no mean definitive) that women authors are much less likely to collaborate internationally; or, alternatively, that available international collaborators tend to be men rather than women, forcing a choice between collaborating with other women or collaborating internationally.

Table VI Centrality and integration of women in CCARD North-South co-publications and selected South publications, 2005–2019

	Thematic set - overall			North-south co-pubs			LMIC publications			LIC publications		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of publications	65,007	117,378	202,883	17,472	32,495	64,040	22,204	40,297	74,735	2,397	4,321	8,060
Share of women amongst authors	29%	31%	33%	26%	29%	31%	26%	28%	30%	25%	27%	30%
Growth index publications with 2+ women authors		1.19	1.13		1.28	1.15		1.23	1.21		1.36	1.17
Network centrality and positions												
Relative difference in centrality degree, women authors against men authors	-37.9%	-37.7%	-28.9%	-39.0%	-39.6%	-29.0%	-42.7%	-40.3%	-36.2%	-65.8%	-14.3%	-21.8%
Share of highly central women authors	7.3%	8.0%	8.0%	7.4%	7.3%	8.4%	7.5%	7.4%	7.2%	5.2%	8.9%	8.6%
Share of highly central men authors	11.2%	11.0%	11.2%	11.1%	11.1%	11.0%	10.7%	10.8%	11.1%	11.5%	10.3%	10.2%
Women-Men integration in co-authorship												
Heterophily women authors - men authors	0.90	0.92	0.91	0.93	0.95	0.93	0.93	0.95	0.92	0.93	0.98	0.98
Homophily women authors - women authors	1.26	1.20	1.18	1.19	1.14	1.16	1.21	1.13	1.18	1.18	1.06	1.05
Homophily men authors - men authors	1.04	1.03	1.04	1.02	1.02	1.03	1.02	1.02	1.03	1.03	1.01	1.01

Note: LMIC: low -middle income countries. LIC: low income countries.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

Table VII Shares of international co-publications by intensity of women authorship in CCARD publications, 2005–2019

	Data set - overall			High women authorship			Mid women authorship			Low women authorship		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of papers	65,000	117,366	202,857	5,050	8,835	14,730	25,686	52,763	101,168	32,523	53,517	83,603
Collaboration practices												
Share international co-publications	33.1%	34.2%	38.2%	14.7%	15.5%	18.1%	36.2%	37.3%	40.5%	34.9%	35.4%	40.3%
Share North-South co-publications	27.8%	28.2%	31.8%	12.9%	12.7%	14.6%	30.3%	31.0%	33.8%	29.3%	29.0%	33.4%
Growth index North-South co-publications		1.01	1.13		0.99	1.15		1.02	1.09		0.99	1.15
Share South-South co-publications	2.1%	2.6%	3.3%	1.1%	1.2%	1.4%	2.1%	2.5%	3.2%	2.4%	2.9%	3.9%
Share North-North co-publications	3.7%	3.8%	3.3%	1.5%	2.2%	2.4%	4.4%	4.2%	3.7%	3.7%	3.8%	3.1%

Note: High women participation publications: 75% of more of women authorship. Mid women participation publications: 75 to 25% of women authorship.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

Key finding 7: It appears likely that North-South co-publications did not meaningfully impact positions for women in the network. North-South co-publications recorded slightly less women authorships than in the overall field (31% to 33% in the 2015–2019 period). Differences in average centrality between women and men did not differ meaningfully in those publications. The gap in average centrality for women in North-South co-publications has decreased by roughly the same magnitude as in the main dataset (from -40% to -29% between 2010–2014 and 2015–2019).

It can also be noted that publications with a high proportion of women authorships (75% or more) recorded much lower values for international co-publications (18% against 38% in the main dataset) and for North-South co-publications (15% against 32%). This observation might provide a signal (by no means definitive) that women authors are less likely to collaborate internationally; or, alternatively, that the available international collaborators tend to be men rather than women, creating a potential trade-off between collaborating with other women or collaborating internationally.

2.6 Have North-South co-publications impacted network positions for southern actors?

On first thought, North-South collaborations may be considered as useful vehicles for increasing scientific capacity and participation of Global South countries. Some authors, however, have raised the possibility that South-based authors are predominantly integrated in international teams holding subordinate roles,[5] therefore contradicting some of the objectives of these initiatives.

Within CCARD North-South co-publications, the average betweenness centrality of South authors was at a wider gap to the average centrality of North authors, than in the CCARD field overall (Table VIII). In the recent period, the centrality gap was of -44% for South authors, while it was of -33% in the full CCARD set. Differences in centrality gaps were even more pronounced for shares of North and South authors amongst the most central authors in the field. For instance, 8 % of South authors were amongst the most central authors in North-South co-publications, against 13% of North authors. These gaps have also increased in time, but not at a faster pace than what was also observed in the full thematic set.

Startingly, centrality gaps transform into leads for South authors in LMIC and LIC publications. While it is true that these breakdowns may be expected to be characterized by higher numbers of South authors than North authors, this was in fact only true for LMIC papers (67% of South authors, data not shown – against 63% in the full thematic set). LIC publications recorded only 50% of South authors, and nevertheless recorded strong leads on centrality for South authors.

LIC publications saw South author lead on centrality to North authors recorded at 29% in the recent period. This lead was down from a peak of 70% in the prior period. Findings on shares of South authors that were highly central authors displayed great leads. For instance, 12% of South authors were in the top decile for centrality amongst LIC publications, against almost 9% for North authors.

[5] Mosbah-Natanson, S., & Gingras, Y. (2014). The globalization of social sciences? Evidence from a quantitative analysis of 30 years of production, collaboration and citations in the social sciences (1980-2009). *Current Sociology*, *62*(5), pp. 626–646. doi:10.1177/0011392113498866.

Table VIII Centrality and integration of South authors in CCARD North-South co-publications and selected South publications, 2005–2019

	Thematic set - overall			North-south co-pubs			LMIC publications			LIC publications		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of papers	65,007	117,378	202,883	17,472	32,495	64,040	22,204	40,297	74,735	2,397	4,321	8,060
Collaboration practices												
Share international co-publications	33.1%	34.2%	38.2%	100.0%	100.0%	100.0%	37.2%	40.1%	44.9%	77.1%	80.3%	80.2%
Share North-South co-publications	27.2%	27.8%	31.6%	99.9%	99.9%	99.9%	33.5%	35.6%	39.4%	65.0%	68.3%	68.2%
Growth index North-South co-publications		1.03	1.14		1.00	1.00		1.07	1.11		1.05	1.00
Share South-South co-publications	2.1%	2.6%	3.3%	0.1%	0.1%	0.1%	3.7%	4.5%	5.5%	12.1%	12.0%	11.9%
Share North-North co-publications	3.7%	3.8%	3.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Network centrality and positions												
Relative difference in centrality degree, South authors against North authors	-20.1%	-23.6%	-33.2%	-34.3%	-40.1%	-44.3%	19.8%	10.9%	3.4%	57.8%	70.3%	29.3%
Share of highly central North authors	11.1%	11.6%	12.4%	12.0%	12.8%	13.4%	9.5%	10.0%	9.5%	7.7%	8.3%	8.6%
Share of highly central South authors	9.5%	9.2%	8.9%	8.2%	7.6%	7.6%	10.3%	10.2%	10.6%	12.9%	12.4%	12.1%
North-South integration in co-authorship												
Heterophily North authors - South authors	1.23	1.14	1.08	1.39	1.26	1.18	1.23	1.14	1.07	1.16	1.07	1.02
Homophily North authors - North authors	0.84	0.91	0.95	0.71	0.82	0.89	0.78	0.88	0.94	0.79	0.91	0.98
Homophily South authors - South authors	0.67	0.79	0.88	0.48	0.61	0.73	0.76	0.85	0.92	0.88	0.94	0.98

Note: LMIC: low -middle income countries. LIC: low income countries.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

LMIC publications also saw leads in centrality for South authors, although moderate ones. By the recent period this lead had been reduce to 3%, but it reached almost 20% in the initial period. South and North authors were roughly on par for the proportion of their groups falling within the top decile of centrality.

LMIC and LIC publications have both seen the leads of their South authors on average centrality decrease over time, however. The findings available do not provide any clear indication as to what might have driven decreases here, and the panel of potential factors to consider is large – this therefore needs to remain a consideration for future work.

2.7 Has OA publishing impacted network positions for women actors?

Within the CCARD set, Open Access (OA) publishing might have had a minor positive impact on women's average centrality and integration in co-authorship, or those women publishing most in OA might not be representative of women in the full network (they would be more productive and central – do note however that proportionally and absolutely, there were more women authors in the network of OA publications than in the network of non-OA publications). Average centrality in the group of women authors within OA publications was -25% below that of men within OA publications; this compared to a difference of -31% in the group of non-OA publications (Table IX). The shares of women falling within highly central authors was roughly the same in both breakdowns, and similar to measurement in the full publication set, however. Women-men heterophily was very slightly above for OA publications (0.92 in the latest period in OA publications, against 0.88 in non-OA publications). This difference was mostly driven by a decrease in women homophily in the OA publication set (1.16 compared to 1.27 in the non-OA set for 2015–2019). Changes in time for all these dimensions did not appear to diverge meaningfully in the OA and non-OA sets from the trend recorded in the full CCARD field.

From a high-level view, OA publishing did not appear correlated with strong increases in women centrality and integration, although it might have made a positive (if small) contribution. Further modelling efforts would be needed to confirm this.

Table IX Centrality and integration of women in CCARD OA publications, 2005–2019

	Thematic set - overall			OA papers			non-OA papers		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics									
Number of publications	65,007	117,378	202,883	24,557	48,747	72,529	33,705	57,248	108,233
Share of women amongst authors	29%	31%	33%	29%	32%	34%	30%	31%	32%
Growth index publications with 2+ women authors		1.19	1.13		1.22	1.16		1.15	1.14
Network centrality and positions									
Relative difference in centrality degree, women authors against men authors	-37.9%	-37.7%	-28.9%	-33.4%	-37.6%	-25.4%	-43.1%	-35.3%	-31.3%
Share of highly central women authors	7.3%	8.0%	8.0%	8.2%	7.4%	8.4%	7.6%	8.1%	8.0%
Share of highly central men authors	11.2%	11.0%	11.2%	11.0%	11.2%	11.0%	10.7%	10.9%	11.1%
Women-Men integration in co-authorship									
Heterophily women authors - men authors	0.90	0.92	0.91	0.92	0.93	0.92	0.87	0.89	0.88
Homophily women authors - women authors	1.26	1.20	1.18	1.20	1.17	1.16	1.33	1.27	1.27
Homophily men authors - men authors	1.04	1.03	1.04	1.03	1.03	1.04	1.05	1.05	1.05

Source: Prepared by Science-Metrix using Scopus (Elsevier) data

Key finding 8: Differences on the centrality of women within OA publications compared to other publications were both positive and negative, depending on the individual indicator. Women have made gains in average centrality in OA publications over time (from a difference of-38% in 2010–2014 to -25% in 2015–2019), replicating the trends found in the main dataset and in non-OA publications. OA publications did see higher women average centrality than non-OA publications in the recent period, but this pattern was not stable in time.

2.8 Has OA publishing impacted network positions for South actors?

For researchers from a Global South country, at least in the CCARD field, OA publishing may have had negative ramifications for centrality and integration in co-authorship (Table X). While OA CCARD publications recorded higher levels of international co-publications (46%) and North-South co-publications (32%) than their non-OA counterparts, OA publications saw a much wider gap for South authors in centrality (-38%) than non-OA publications (-22%; both findings from the recent period). Centrality gap between North and South actors by OA status did not change between the initial and third period (centrality gaps were reduced in the second period).

Heterophily was also lower in OA publications than it was in non-OA publications (1.01 to 1.21 in the recent period). Both scores have seen decreases in time, like in the full CCARD set.

Table X Centrality and integration of South authors in CCARD OA publications, 2005–2019

	Thematic set - overall			OA papers			non-OA papers		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics									
Number of papers	65,007	117,378	202,883	24,557	48,747	72,529	33,705	57,248	108,233
Collaboration practices									
Share international co-publications	33.1%	34.2%	38.2%	37.2%	38.4%	46.3%	31.3%	32.3%	36.5%
Share North-South co-publications	27.2%	27.8%	31.6%	30.3%	30.9%	38.1%	25.9%	26.5%	30.4%
Growth index North-South co-publications		1.03	1.14		1.03	1.24		1.03	1.15
Share South-South co-publications	2.1%	2.6%	3.3%	2.3%	2.7%	3.5%	2.2%	2.5%	3.3%
Share North-North co-publications	3.7%	3.8%	3.3%	4.6%	4.8%	4.6%	3.3%	3.2%	2.8%
Network centrality and positions									
Relative difference in centrality degree, South authors against North authors	-20.1%	-23.6%	-33.2%	-39.1%	-28.4%	-37.6%	-22.1%	-9.6%	-21.9%
Share of highly central North authors	11.1%	11.6%	12.4%	12.1%	12.0%	12.8%	10.5%	10.2%	11.1%
Share of highly central South authors	9.5%	9.2%	8.9%	8.7%	8.9%	8.4%	9.8%	10.0%	9.6%
North-South integration in co-authorship									
Heterophily North authors - South authors	1.23	1.14	1.08	1.15	1.07	1.01	1.33	1.25	1.21
Homophily North authors - North authors	0.84	0.91	0.95	0.90	0.96	0.99	0.75	0.81	0.84
Homophily South authors - South authors	0.67	0.79	0.88	0.76	0.89	0.98	0.58	0.66	0.71

Source: Prepared by Science-Metrix using Scopus (Elsevier) data

Given how OA has been defended as a movement for change in research practices that support the integration of researchers from lower-income countries in research systems, this result might appear counter-intuitive. Others, however, have already reported on similar findings, showing how OA may in

fact reinforce existing hierarchies in research status.[6] Siler and colleagues indeed report findings that policies for OA fee waiving towards authors from lower income countries are not systematically applied. Science-Metrix hypothesizes that this may particularly be the case for North-South co-publications.

Key finding 9: OA publishing may be associated with lower levels of average centrality for South authors, although this dimension did not appear to affect temporal trends in South author centrality. OA CCARD publications recorded higher levels of international co-publications (46%) and North-South co-publications (32%) than their non-OA counterparts. Yet, OA publications saw a much wider gap for South authors in centrality (-38%) than non-OA publications (-22%) in the recent period. OA publication replicated the trend already found in the overall dataset of decreasing relative average centrality for South authors.

2.9 What has been the contribution of IDRC and FCDO-supported publications to these developments?

From the list of publications provided to Science-Metrix, it is possible to broadly characterize the type of research that has been supported by the IDRC and FCDO programs of interest, and their achievements on some dimensions of interest (Table XI). Note that it is not possible to robustly measure centrality and heterophily in these publications, because doing so would have entailed considering the set of these publications has a single, unitary co-authorship network – a highly artificial situation.

Table XI Collaborative, gender equity and OA achievements of IDRC-FCDO publications, 2005–2019

	Thematic set - overall			IDRC-FCDO CCARD pubs		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics						
Number of papers	65,007	117,378	202,883	1	42	403
International collaboration practices						
Share international co-publications	33.1%	34.2%	38.2%		66.7%	70.7%
Share North-South co-publications	27.2%	27.8%	31.6%		47.6%	60.0%
Growth index North-South co-publications		1.03	1.14			1.26
Share South-South co-publications	2.1%	2.6%	3.3%		11.9%	3.7%
Share North-North co-publications	3.7%	3.8%	3.3%		7.1%	6.9%
Women authorship						
Share Women authorship	29.1%	31.3%	33.1%		34.4%	34.4%
Cross-disciplinarity						
Share of highly multidisciplinary papers	11.8%	13.5%	15.8%		23.5%	29.6%
Share of highly interdisciplinary papers	12.8%	14.2%	15.6%		24.1%	28.3%
OA publishing						
Share OA papers	42.1%	46.0%	40.1%		40.5%	66.8%

Source: Prepared by Science-Metrix using Scopus (Elsevier) data

[6] Siler, K., Haustein, S., Smith, E., Larivière, V., & Alperin, J. P. (2018). Authorial and institutional stratification in open access publishing: the case of global health research. *PeerJ*, *6*, p. e4269. doi:10.7717/peerj.4269.

These findings confirmed that the international collaborative element in IDRC-FCDO programs was successfully implemented within supported publications. A share of 71% of these publications were written as international co-publications, and 60% were North-South co-publications.

Shares of authorships held by women in IDRC-FCDO publications were only slightly above the figure found for the whole CCARD field, at 34% (against 31%-33%).

Exceptionally high proportions of IDRC-FCDO were highly cross-disciplinary. A share of 30% of these publications were found to be amongst the top decile of publications for multidisciplinarity in the recent period, two times more (200%) than the expected figure. Again, in the recent period 28% of IDRC-FCDO publications fell within the most interdisciplinary publications. These shares compared against equivalent measurements of 16% in the full CCARD field, which in themselves already indicate that CCARD was already a highly cross-disciplinary field to start with.

While this was not so much the case in the 2010–2014 period, the 2015–2019 period saw high levels of OA publishing for IDRC-FCDO supported articles (67%) relative to the full CCARD set (40%).

Key finding 10: Large portions of the 403 IDRC-FCDO articles published between 2015 and 2019 contained highly collaborative, cross-disciplinary and open research. A share of 60% of those publications were North-South co-publications. Almost 30% of these publications were either amongst the most highly multidisciplinary and/or interdisciplinary in their fields (or put differently, three times the reference figure of 10%). A 67% share of these publications were made available under various OA modalities. IDRC-FCDO funding did not appear to have increased shares of authorships by women, however. A 34% share of IDRC-FCDO publications' authorships were held by women, very close to the overall figure found in the field under study – although it would require a more sophisticated *evaluation* design (as opposed to the network analysis conducted here) to evaluate robustly the specific outcomes enabled by IDRC-FCDO.

2.10 Highlights from a smaller research front: Scopus Topic Cluster on climate change adaptation, resilience, and disasters research

To investigate the influence of levels of aggregation on the findings presented so far, network analyses were also produced for a different specification of the CCARD field, one that used the Scopus topics of prominence product. Scopus Topic Cluster #438 combines topics of prominence that are related to climate change adaptation; resilience; but also natural disaster management (see an overview of the cluster composition in Table XII). This clustering of publications into topics and one Topic Cluster was produced by Scopus teams on the basis of citation networks between publications, with labelling provided by a combination of text mining and manual curation. Cluster 438 is much smaller than the CCARD publication set, with close to 30,000 publications recorded between 2005–2019 rather than 385,000 articles in the second set. It also overlaps with only 80 of the IDRC-FCDO publications, against 403 in the CCARD publication set. This section will provide only key results from this analysis and will highlight differences from what has been found so far. It should be noted, furthermore, that the definition of Cluster 438 did not involve filtering out for publications by North authors only that did not include a South-based component.

Table XII Most voluminous Scopus Topics of prominence in Topic Cluster 438

Scopus topic ID	Keywords	Pubs 2005-2019
1567	Adaptation; Adaptation planning; Climate change	4,493
3646	Governance; Resilience; Resilience thinking	3,448
10049	Flood; Flood loss; Floods	1,963
8720	Disaster; Disasters; Emergency managers	1,580
16054	Disasters; Resilience; Resilience assessment	1,515
18319	Disaster; Disasters; Post-disaster housing	1,327
13602	Flood; Flood insurance; Risk	1,181
12561	Climate change; Environmental migration; Migration	980
33716	Disaster; Natural disaster; Synthetic control	552
35386	Adaptation; Climate change; Water supply	428
27217	Bangladesh; Disaster; Natural hazards	384
28076	Arch dams; Dam; Dams	375
30758	Disaster; Flood; Flood disaster	375
35610	Concrete dams; Dam deformation; Dams	367
45376	Alarm systems; Disasters; Natural disasters	337
39724	Flood; Risk; Technical networks	305
29122	Hazard mitigation; Plan; Planning	301
36265	Flood control; Flood season; Floods	264
46053	Community; Disaster resilience; Disasters	262
44837	Disasters; Natural disasters; Risk management	260
44478	Disaster; Disaster preparedness; Disasters	258
42536	Forecasting; Neural networks; Sample generation	254
38355	Flood control; Floods; Waterlogging	230
44475	EU; River basin; Water	208
42649	Disaster; Disasters; Typhoon disaster	205
51267	Coast; Coastal vulnerability; Vulnerability	195
51969	Disaster; Disasters; Natural disasters	186
49533	Disasters; Management; Publications Switzerland	173
52002	Groundwater; River basin; Water	171
61482	Disasters; Power supply; Telecommunication networks	170
48204	Disaster; Special issue; Woman	161
56007	Disasters; Emergency plans; Management	159
69771	Coastal zones; Land subsidence; Subsidence	159
50843	Built environment; Disaster; Disasters	148
47347	Annual rainfall; Drought; Rainfall	147
53557	Flood; Flood protection; Risk management	140
40432	Disaster; Disasters; Flood events	139
74748	Disaster; Disasters; Mindanao Philippines	132
67937	Disaster; Disaster waste; Disasters	129
55877	Austrian economics; Culture; Disaster	128

Note: Top 40 Topics based on publication count between 2005 and 2019. Remaining 81 Topics were associated with 22 to 125 publications each over the period.

2.10.1Women centrality and co-authorship integration in Cluster 438

Looking at women authors' centrality and heterophily in this network, it was found that they held a lead on average centrality to men in the most recent period (7% ahead) (Table XIII). Correspondingly, slightly more women fell into the top decile of most central authors (12%) than men did (10%). These figures have remained stable in time, while observations for average centrality showed more equivalent achievements in the two prior periods. Men-women heterophily was very close to what was observed in the CCARD set, slightly below expectations and driven by women homophily.

We now turn to how North-South co-publications, multidisciplinary and interdisciplinary co-publications might be correlated to observed differences in the centrality and integration of women and men in Cluster 438 (Table XIII). In North-South co-publications (which correspond to a smaller share of Cluster 438 than they did in the CCARD field, for methodological reasons), women's average centrality relative to men was somewhat more volatile, standing at a gap of -6% in 2015–2019 but 19% ahead in the prior period. Findings in terms of shares of women and men amongst the most central authors, and heterophily findings, did not meaningfully deviate from what is seen in Cluster 438 overall. Given these findings, it was not possible to determine whether North-South co-publications might have influenced the relative position of women in Cluster 438.

By contrast, findings on the multidisciplinarity dimension showed that such publications were likely correlated to improvement in women's positions in this network. In the 2015–2019 period, 15% of women were amongst the most central authors within highly multidisciplinary publications, up from 10% and 12% in prior periods. Average betweenness centrality was 42% higher for women than men in 2015–2019. Centrality for women within the network of highly interdisciplinary publications was also higher than men's, although at its highest in the very first period considered (2005-2009). These results point towards a positive correlation of cross-disciplinary articles on women centrality, although it is not clear whether it is the cross-disciplinary nature of research that explain observed differences or differences in the type of women/men researchers participating in these modes of research.

2.10.2 South author centrality and co-authorship integration in Cluster 438

The analysis of Cluster 438 recorded large gaps on centrality of South authors to North authors, although these gaps may be greater in this component of the analysis by virtue of including all north authors (Table XIV). In this case, North-South and multidisciplinary co-publications do appear to correlate positively with the positions of South authors, although it is not clear whether these correlations are stable or increasing in time, especially when comparing average centrality and share of highly central authors.

By the 2015–2019 period, 26% of Cluster 438 publications were international co-publications. North-North co-publications made the largest part of these co-publications, accounting for 15% of all Cluster 438 publications. North-South co-publications accounted for 10% of all Cluster 438 publications, South-South for only 1%.

Table XIII Centrality and integration of women in Cluster 438 North-South and cross-disciplinary publications, 2005–2019

	Topic 438 overall			North-south co-pubs			HMP10%			HIP10%		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of publications	5,168	9,258	15,326	218	560	1,583	438	955	2,016	618	1,698	3,892
Share of women amongst authors	29.8%	32.9%	37.5%				29.1%	31.9%	38.2%	35%	35%	40%
Growth index publications with 2+ women authors		1.20	1.41					1.22	1.34		0.97	1.28
Network centrality and positions												
Relative difference in centrality degree, women authors against men authors	-3.5%	-1.1%	6.7%	-12.2%	18.8%	-6.4%	106.7%	-24.8%	41.9%	90.5%	4.8%	19.3%
Share of highly central women authors	11.8%	11.7%	11.8%	11.1%	10.3%	11.2%	11.8%	10.1%	14.7%	15.6%	10.5%	12.4%
Share of highly central men authors	10.1%	10.1%	10.3%	9.7%	9.7%	10.2%	10.9%	10.5%	9.2%	9.2%	10.9%	9.8%
Women-Men integration in co-authorship												
Heterophily women authors - men authors	0.87	0.88	0.90	0.91	0.91	0.92	0.94	0.88	0.91	0.86	0.91	0.89
Homophily women authors - women authors	1.38	1.29	1.18	1.32	1.23	1.12	1.18	1.27	1.13	1.32	1.20	1.16
Homophily men authors - men authors	1.04	1.05	1.06	1.02	1.04	1.05	1.02	1.05	1.06	1.06	1.05	1.07

Note: $HMP_{10\%}$: Share of highly (top decile) multidisciplinary publications; $HIP_{10\%}$: Share of highly (top decile) interdisciplinary publications.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

Table XIV Centrality and integration of South authors in Cluster 438 North-South and cross-disciplinary publications, 2005–2019

	Thematic set - overall			North-south co-pubs			$HMP_{10\%}$			$HIP_{10\%}$		
	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19	2005-09	2010-14	2015-19
Descriptive statistics												
Number of papers	5,168	9,258	15,326	218	560	1,583	438	955	2,016	618	1,698	3,892
Collaboration practices												
Share international co-publications	14.7%	18.6%	26.4%				32.3%	33.4%	43.0%	22.9%	24.2%	31.9%
Share North-South co-publications	4.3%	6.1%	10.4%				10.1%	11.6%	20.7%	6.5%	7.5%	13.4%
Growth index North-South co-publications		1.43	1.71		n/a			1.15	1.79		1.15	1.80
Share South-South co-publications	0.2%	0.5%	0.8%				0.2%	0.4%	1.2%	0.5%	0.6%	0.7%
Share North-North co-publications	10.2%	12.0%	15.2%				21.9%	21.4%	21.0%	15.9%	16.2%	17.8%
Network centrality and positions												
Relative difference in centrality degree, South authors against North authors	-73.3%	-84.2%	-63.2%	46.0%	-47.5%	-15.8%	-93.2%	-89.8%	-33.0%	-99.5%	-85.7%	-68.1%
Share of highly central North authors	10.4%	11.9%	12.0%	8.9%	11.2%	10.4%	10.5%	11.2%	11.3%	10.9%	11.0%	11.7%
Share of highly central South authors	7.2%	2.1%	4.8%	13.2%	8.2%	9.1%	4.8%	3.0%	6.3%	0.0%	3.3%	4.5%
North-South integration in co-authorship												
Heterophily North authors - South authors	0.87	0.86	0.88	1.42	1.31	1.23	0.72	0.87	0.83	0.89	0.88	0.93
Homophily North authors - North authors	1.03	1.03	1.03	0.65	0.78	0.85	1.07	1.03	1.05	1.02	1.03	1.02
Homophily South authors - South authors	1.70	1.56	1.46	0.49	0.57	0.65	2.11	1.62	1.56	1.63	1.53	1.27

Note: $HMP_{10\%}$: Share of highly (top decile) multidisciplinary publications; $HIP_{10\%}$: Share of highly (top decile) interdisciplinary publications.
Source: Prepared by Science-Metrix using Scopus (Elsevier) data

The move to the Cluster (and the lack of a Global-South oriented research filter) has negative repercussions on centrality measures for South authors. In this breakdown, they lag by -63% behind North authors in average centrality. When the same Global South filter as applied in the CCARD dataset, this lag was reduced to -34%, close to the gap observed in the CCARD publication set. Unsurprisingly given the observation just made, there is a large difference in the shares of South and North authors that fell within the top decile for author centrality (5% for the first group, 12% for the second). Centrality measures do not fall within a single temporal trend, and it was not possible to reach a reliable conclusion on this dimension.

In the overall Cluster 438 publication set, less heterophily than expected was recorded (0.88 in the latest period). This measurement appeared driven by homophily amongst South authors (1.46, again in 2015–2019). These observations here are very likely to capture the impact of the inclusion of a criterion for South-based work in the CCARD publication set. They show that the CCARD publication set's inclusion criteria skewed the resulting distribution towards northern authors with strong profiles in terms of North-South collaborations. This last measure considerably went down in the subset of Cluster 438 North-South co-publications, which is examined next.

Again, North-South co-publications accounted for 10% of Cluster 438 publications. However, in this breakdown, South authors have tended to be in much more advantageous network positions. South authors were at a lead to North authors in centrality in the 2005–2009 period (by 46% on average centrality; 13% of the group falling within the top decile of author centrality, against 9% of North authors). Their lead turned to a gap in 2010–2014 (-48%). The gap remained in the recent period but was greatly reduced (-16% on average centrality; 9% of South authors in the top decile for centrality, against 10% for North authors).

In terms of integration in co-authorship, North-South co-publications expectedly raised heterophily above expectations (1.23 in the recent period).

The two subnetworks of cross-disciplinarity publications saw higher shares of international co-publications, across all categories considered here. This was particularly true of highly multidisciplinary publications, which, to take just one example, saw 21% of North-South co-publications. Despite this, the subnetwork of interdisciplinary co-publications recorded centrality observations that were roughly on par with those of Cluster 438 overall (except in the first analytical period, where South authors in interdisciplinary papers were at very large gaps to North authors). Centrality measurements in multidisciplinary papers were not more advantageous to South authors in the first two periods, but they greatly improved recently. The gap on average centrality had decreased to -33%, notably. While the share of South authors amongst the top decile for centrality was quite low here (6%), it has improved in time. It is not possible at this stage, however, to definitively link this improvement to cross-disciplinary practices.

3 Conclusions and suggestions for further work

Taken together, the findings presented above convey a picture of "stratified integration" in the field of 'climate adaptation and resilience for development' (CCARD) field. Women and Southern authors participated in this field (although not as numerous as Northern or men authors) and they were included in collaborative opportunities. Yet their individual centrality scores were, on average, not as high as those of their northern or male counterparts.

3.1 High-level summary of findings on authors' centrality

Women authors in CCARD research collectively recorded average betweenness centrality that were -29% below those of men, between 2015 and 2019. The gap has reduced over time, yet this trend of a reduction appeared as much in the overall data set as in the subsets of cross-disciplinary, North-South and Open Access (OA) publications. Cross-disciplinary publications appear to positively correlate with women's centrality scores. OA publishing was also positively correlated with this dimension, albeit with a modest effect. North-South co-publication did not correlate with women authors' centrality, neither positively nor negatively.

Southern authors were collectively, on average, at of gap of -33% below Northern authors for their centrality scores in the CCARD set (for the period 2015-2019). This gap has widened over time. Cross-disciplinary co-publications, North-South co-publications and OA publications were each associated with even greater gaps when Southern author centrality is compared to that of Northern authors. The trend towards a deterioration in time of southern authors' centrality relative to that of northern authors was found in both the overall CCARD set and the subsets of cross-disciplinary, North-South and OA publications.

3.2 Counter-observations and nuances

Despite observations of potential negative impacts of cross-disciplinary and North-South collaboration on South author centrality in the CCARD set, findings for the ancillary analysis of Cluster 438 provide a somewhat different picture. In cross-disciplinary and North-South co-publications, the relative position of South authors to North authors improved quite notably as compared to the full 438 set. While it is not possible here to explain why this is the case, one hypothesis is that cross-disciplinary and North-South co-publications mitigate the starkest gaps in North-South centrality (given the inclusion of all North authors in the analysis of Cluster 438). However, it is also possible that this difference in scores was driven by the differences in methods for building the publication sets themselves.

It must also be noted that publications from LMIC and particularly LIC countries often diverged meaningfully from North-South co-publications and CCARD publications overall in terms of South author centrality. It may well be that the impacts of cross-disciplinary and North-South collaborative publications may be highly differentiated depending on the exact location where they are realized, or at least on the income and development status of these locations.

3.3 Core limitation

Centrality and heterophily measures presented in the analytical breakdowns used in this study do not represent the contribution of the analytical groups to the overall network; but calculated separately for subnetworks composed of publications falling in a given breakdown. Therefore, it cannot be established with full certainty that the differences we see in the breakdown sets are responsible for the trends in the overall set.

Results on the outcomes of cross-disciplinary and North-South collaborations for women and South author centrality and integration are also limited by the inability to perform controls for confounding factors within the scope of the current study. For instance, it is known that women tend to publish less than men; and that women tend to conduct more interdisciplinary research. These two factors might help explain why women's average centrality lagged behind that of men in our findings, but it could also be that one of these two factors is in fact fully determined by the other. Women's lower centrality scores might also be explained by lower average productivity, because prior findings raise the possibility that interdisciplinary research in and of itself may not lead to lower centrality but rather results in lower productivity. If the above is true, the question would then be whether it is the higher propensity towards interdisciplinarity itself, or the lower productivity, that led to the observed pattern in the North-South centrality gap. In other words, what is the "true" causal factor. Obtaining clearer understanding of these complex causal relationships would require the use of advance statistical modeling techniques.

3.4 Next steps

3.4.1 Limitations to centrality- and network-based studies

Co-authorship based centrality is an indicator that captures the capacity of an investigator to mobilize a network of partners in the conduct of her or his research projects. As most highly successful researchers are arguably successful collaborators and networkers, and that collaborative research generally tends to have broader readership as captured by citation levels,[7] it can be argued that average author-level centrality in a group of interest is an appropriate measurement for capturing improvements in southern and women's authors status and conditions within the global science system.

Nevertheless, arguments could also be made against the use of centrality as an indicator of international scientific development. Foremost, is an oft-made observation that, statistically, centrality tends to be most strongly correlated with output volume.[8] Authors with the highest volumes of articles (and, relatedly, seniority defined as total active years of research publishing) tend to hold the highest author-level centralities. Following this logic, the best intervention to increase centrality would be to support higher publication volumes. Yet, this intervention may be highly difficult to implement, given how publication volume may be directly related to the availability of costly research infrastructures in a given country. Put

[7] Struck, D. B., Roberge, G., & Campbell, D. (2017). The influence of open access, gender and co-authorship on citation scores. *Science, Technology & Innovation (STI) Indicators 2017*. Retrieved from https://sti2017.ifris.org/wp-content/uploads/2017/11/oa-db-struck-et-al.pdf.

[8] Gui, Q., Liu, C., & Du, D. (2018). Does network position foster knowledge production? Evidence from international scientific collaboration network. *Growth and Change*, *49*(4), pp. 594–611. doi:10.1111/grow.12263.

shortly, it might be that increasing southern or women centrality can only be achieved with sustained and cross-cutting investments in research infrastructures, rather than with discrete programs with a set duration. Science-Metrix, however, does not posses the international development expertise to provide a definitive answer to this question.

If IDRC possesses robust other evidence that projects with: i) a high intensity of women participations; ii) with a high intensity of southern author participation, or North-South collaboration; or iii) that are highly cross-disciplinary, help the organization achieve its goals, then the direct measure of these dimensions may be a cost-effective way of characterizing fields or programs of interest. Program evaluation methods revolving around these three indicators may also be more apt to recover the micro- or meso-scale outcomes from funding programs than centrality-based indicators.

3.4.2 Design possibilities and limitations for a regression analysis of women or southern author centrality

The results obtained so far provide initial signals on how multi-country collaboration, multidisciplinarity and OA practices have supported South and women author centrality. Science-Metrix proposes to conduct further work in the form of a regression analysis focusing on one main outcome variable of interest and a small selection of explanatory variables, based on the signals already obtained. Of course, the analysis proposed controls for the confounding factors and other limitations already identified, although Science-Metrix cannot guarantee it will produce statistically significant results. In addition, it remains a centrality-based analysis, and therefore will remain limited by the factors identified.

We suggest that the regression analysis should include one or more independent variable(s) that are direct targets of intervention in IDRC-FCDO programs. For instance, probing the influence of "North-South co-publications" (core explanatory variable) on "South author centrality" or "women author centrality" (outcome variables) may be of direct relevance to IDRC, if IDRC grant mechanisms have formal requirements to support North-South co-publications. To take a hypothetical example, if OA publishing is enforced by IDRC for its grantees, but IDRC does not control for diversity in disciplinarity backgrounds of its funded teams, then probing the "influence of OA publications on South author centrality" may have more program-level implications than probing the "influence of multidisciplinary publications on South author centrality".

Below is a high-level design for a potential regression analysis. As this is illustrative, Science-Metrix cannot guarantee at this stage that any of the variables presented will make final inclusion in the model:

Outcome variable:

- Average centrality: author-level, based on co-publication links

Core model specification:

- Model for women authors
- Model for southern authors

Main explanatory variables

- Share of co-authors in author' publications that are women (author-level women-men heterophily)
- Share of author's publications that are women-men co-publications
- Share of co-authors in author' publications that are southern (author-level southern-northern heterophily)
- Share of author's publications that are North-South collaborations
- Overall interdisciplinarity score of papers by the author
- Overall multidisciplinarity score of papers by the author

Additional explanatory variables and dummies that can include other dimensions that were of interest in the present study, or confounding factors that need to be tested for:

- Share of publications as first, senior or corresponding author
- Share of authors' publications supported by IDRC-FCDO
- Share of IDRC-FCDO's total publications to which the author has contributed
- Share of OA publications by the author
- Number of publishing years (as a proxy for seniority)
- Number of articles by the author in the CCARD subfield
- Number of articles by the author in all subfields
- Main country of author
- Single most common Topic of prominence of authors' publications

Contribution from each of these variables to the variance on author-level centrality can be computed, with a view to determine which variables explain most of the differences observed in a given model.

3.4.3 Design possibilities and limitations for network studies that evaluate programme or project-level outcomes on southern or women centrality

IDRC asked Science-Metrix to comment on the design that would be needed to study of the specific outcomes of IDRC-FCDO support on southern and women author centrality. Such a study is, in principle, feasible, by restricting the model to only IDRC-FCDO publications while still using average author-level centralities measured within the CCARD publications set. In this instance, the model will need to employ strong controls for seniority and other confounding factors. Given the smaller sample to work with in such a scenario, there is, however, a higher risk here that the analysis will not result in statistically significant results.

3.4.4 Design possibilities and limitations for network studies that evaluate country-level southern or women centrality

IDRC may also be interested in identifying countries on which to potentially focus their investments and interventions, on the basis of notable increases of either women or southern author average centrality. Doing so is possible, in principle, using the kind of regression analyses described above to control again for factors such as author productivity, seniority and other factors that could not be controlled for in the present study. Countries might be characterized by their average author-level centrality profiles: lagging, leapfrogging, and so forth.

It should be noted, however, that the aggregate number of publications with participation by at least one LIC-based author in the CCARD publication set was 8,060. This is a small sample to work with in a regression model, and may not lead to statistically reliable results. Single LIC countries will not have enough publications associated with them to be included in such a model and reach statistical significance. Therefore, Science-Metrix expects that such a model specification could only be deployed for some UMIC countries, and the LMIC countries with the largest output volumes.

Science-Metrix suggests that, in preparation for such a study, IDRC first determines whether its intervention preference would be for "consolidating champions" by investing in countries seeing high increases in author-level average centralities; or for "reducing inequalities" by focusing on those countries with decreasing author-level average centralities.

3.4.5 Design possibilities and limitations for network studies that identify thematic areas that contribute to southern or women centrality

Defining or capturing research fronts and topics is arguably one of the most vexing goals for current bibliometric practice; and possibly science policy more largely. Isolating research fronts where average southern or women author centrality has greatly increased in recent years will face the full range of limitations associated with identifying increases in centrality already identified above.

In addition, such a project would face additional limitations related to the specific stem of delineating and defining research fronts or topics themselves. First is the issue of delineation. Cutting-edge research fronts, by definition, are small in volume; and therefore difficult to capture through bibliometric methods, until they become more established. Even for larger research fronts, the adequate scale of aggregation is an open question in every project, requiring exploratory work in the process of clarifying the best scale. Scopus Topic Clusters such as Topic Cluster 438 used in this project amounts to a rather large research front (several thousand publications), but may still offer too small a sample for some regression analyses or econometric models. Next scale classification schemes, such as the Science-Metrix subfields (including categories such as Environmental Engineering or Agricultural Economics & Policy), are likely already too large to detect the dynamics of interest here.

Second is the issue of definition. Even with a robust method for delineating clusters of publications, manual examination of publications in a given topic might provide conclusions about the content of these articles that are quite different from the expected thematic making of the field, obtained, say, by expert assessment. It is well known that researchers tend to latch on some keywords and catch phrases to increase their chances of getting funding, among others.[9] Keyword-based might not yield thematically coherent publication sets, or at least not coherent along the lines initially envisioned by a sponsor. Even using methods such as citation-based clustering, the contours of research fronts tend not to align with expectations and established classifications into institutional disciplines or bibliometric classifications. It is often difficult to reconcile data-driven clustering methods with human expectations. Still today, doing

[9] Morris, N., & Rip, A. (2006). Scientists' coping strategies in an evolving research system: the case of life scientists in the UK. *Science and Public Policy*, *33*(4), pp. 253–263.

so most likely requires scanning and reading of random samples of publications contained in research fronts.

Science-Metrix' suggestion is that such a goal is best tackled with peer review; qualitative assessment; or policy analysis. Bibliometrics could be used to establish an initial 'shortlist' of research fronts of interest, and provide reading lists of publications for further examination.

3.4.6 Best practices in collecting qualitative evidence to support and triangulate with bibliometric studies

Science-Metrix offers the following recommendations to enhance the value of future bibliometric studies for the IDRC:

- Conducting a review of the existing academic, international development and policy literatures on 1) gender inequalities in research systems and practices, 2) gender inequalities in North-South cooperation projects, and 3) particular situations of women researchers from the Global South, might allow to better understand some of the impact of collaborative programs on women researchers; and help formulate efficient follow-up research designs requiring less exploratory work.
- Possibly as a sub-analysis of the work stream suggested immediately above, a mix of desk research, policy analysis and descriptive bibliometrics could be employed to assess Global South countries (or a selection thereof) for their relative level of research capacity. These qualitative findings could be combined with some of the data sets produced as part of this project to assess the extent to which centrality of South authors was determined by productivity levels and differences in scientific infrastructure between countries.
- As part of generic program management activities, IDRC should strive to collect data on the gender and South-North composition of the research teams it supports. Perceptions of women and South-based authors could be collected with a view to obtains self-assessments of the degree of equality between South and North; and women and men authors within those teams. In conducting such assessments, it is important to seek out answers from all team participants, rather than only principal investigators. These data sets should be organized with funded projects as the basic measurement unit. Projects should also be linked to the calls (program name and year) through which they were funded.

4 Appendix A – Methods

4.1 Databases and publication sets

4.1.1 Bibliometric database

For the purpose of this project, Science-Metrix used the Scopus database, produced by Elsevier. Scopus provides comprehensive coverage of the scholarly literature by indexing more than 43 million publications, published in some 50,000 peer-reviewed journals and conference proceedings since 1996. Scopus also provides the names and affiliations of all authors appearing in peer-reviewed publications, making it possible to identify publications produced by individual researchers and the institutions with which they are affiliated.

The document types included in the Scopus analysis are articles, reviews, short surveys and conference proceedings. Unless stated otherwise, the tables and figures deriving from Scopus data include all the aforementioned document types. The version the production database used for this project had complete coverage of articles published up until 2019. This project covered the period 2005–2019.

4.1.2 Scientific publications included in the analysis

The publication set used in this study was described in detail in section 1.2.1.

4.2 Analytical periods

Analytical periods used for this study were 2015-2019; 2010-2014; and 2005-2009.

4.3 Bibliometric indicators

4.3.1 Output and impact indicators

4.3.1.1 Number of publications, using full and fractional counting

This indicator, also called the publication output, is the number of articles published by a given entity (in this case, a researcher); it is calculated using two methods, known as *full counting* and *fractional counting*. Science-Metrix proposes to present results from both methods for this project. Using the full counting method, each entity represented by each author listed on a publication receives full credit (1 publication) for that paper. For example, if a publication is authored by two researchers affiliated to the UK, one to Spain and one to the US, that publication is counted once for the world, once for the UK, once for Spain and once for the US. A publication count based on full counting indicates which entities are involved in the production of an article, regardless of their individual level of contribution to the article. With the fractional counting method, each author on a publication is instead awarded an equal share of the paper so that the sum across all authors equals 1. For example, if a publication is authored by two researchers affiliated to the UK, one to Spain and one to the US, each of the four authors receives a count of 0.25 of a publication, meaning that the UK would receive a count of 0.5 for that publication, with Spain and the US receiving 0.25 each. Although it is not possible to determine the contribution of each author on a

publication, fractional counting provides a more precise representation of the share of each entity compared to the full counting.

4.3.1.2 Share of international co-publications, southern publications or North-South co-publications

The share of international co-publications (or the international co-publication rate, ICR) of a researcher is his or her share of publications that include at least another author from a different country.

Science-Metrix also computed variations of these indicators to capture Global North – Global South co-publications, as well as Global South – Global South co-publications (using the OECD ODA list to assign countries to either category, or another classification scheme to be provided by the IDRC).

4.3.1.3 Share of open access publications

Open access (OA) as a topic in science policy has grown immensely in importance in recent years. Science-Metrix has an intense and long-standing interest in OA and is uniquely well placed to conduct bibliometric research projects on this topic. 1science, Science-Metrix's sister company, has constructed a database of peer-reviewed, OA publications using a web harvester to collect and characterize papers from the web. The definition being applied is a simple one: a publication is OA if it can be accessed for free and without any barrier, such as a subscription or registration.

Because the content of the 1science OA database is cross-referenced to the publications in the Scopus database, Science-Metrix is able to deploy its full arsenal of analysis tools to characterize the research enterprise as it evolves on both sides of the OA divide. For this study, Science-Metrix retrieved the share (%) of thematic journal publications that can be accessed through an OA licence.

4.3.1.4 Identifying authorship by women authors

NamSor™, a European designer of name recognition software, was used in this study to assign gender to publication authors. NamSor™ offers a very high degree of accuracy and recall, in addition to its global scope, covering all languages, alphabets, countries and regions. The NamSor application programming interface (API) is also quite tolerant of typographical errors, a notably useful feature in light of the significant number of input errors in the bibliographic databases.

NamSor have implemented a rigorous protocol to assess the quality of their tool, demonstrating that it can achieve a high recall (i.e., there are very few unknowns) and accuracy (i.e., there are very few false positives) in the United States, Canada, Mexico, Russia, Japan and most European countries. Their validation procedure relies on the use of directories listing names along with their geographic location (i.e., country) and specified gender of titles (Mr. for men and Ms. for women). Using the known gender of individuals, they validate whether their algorithm attributes the correct gender.

For each combination of first name and last name, the API returns a score between -1 and 1. A score of -1 indicates a man with a certainty of 100%, whereas a score of 1 is returned for a woman, again with 100% certainty. A score of 0 denotes that the gender can't be determined at all. In fact, the API does not provide a gender for all values between -0.1 and 0.1.

4.3.2 Crossdisciplinarity indicators

4.3.2.1 Interdisciplinary integration

Examining the material that is cited in a paper offers a reflection of the intellectual content that is being integrated in the underlying research. Accordingly, the integration of material drawn from across disciplinary boundaries is assessed through citation behaviours. The interdisciplinary integration (II) indicator considers (a) the number of different subfields that are being cited, (b) the distribution of those citations across the cited subfields, and (c) the intellectual proximity of those subfields to one another.

For example, a paper that draws on knowledge from four different subfields would have a higher interdisciplinarity score than a paper that draws on only three. Similarly, a paper that cites one subfield 90% of the time and the other subfields only 10% of the time would have a lower score than a paper that cites its various subfields in roughly equal measure. Finally, a paper that integrates knowledge from biology and from chemistry would have a lower score than a paper that integrates knowledge from biology and the performing arts, because the former pair is more intellectually proximate than the latter pair.

4.3.2.2 Highly interdisciplinary papers

For this study, the indicator to be computed shows what share of an entity's papers fall within the top 10% of highly interdisciplinary papers in the world (HIP), with each paper's interdisciplinarity score adjusted to the average of all papers worldwide published in the same subfield and same year (similar to the RC, see above).

4.3.2.3 Multidisciplinary integration

For this study, the index of multidisciplinary integration (MI) relies on Science-Metrix's journal-based classification of science. It reflects the diversity of the prior disciplinary backgrounds of a paper's co-authors. It is computed by adapting the metrics of Porter & Rafols to the disciplinary profile of co-authors in a paper.[10] MI was designed to increase for teams involving authors from different subfields, particularly where these subfields are not frequently connected in Scopus. It is normalized by the paper's subfield and year to avoid coverage biases.[11]

A paper co-authored by authors whose previous papers were distributed across subfields of science in a similar pattern (i.e., having similar relative frequency across subfields) would score lower than a paper bringing together authors with different backgrounds (as measured by the subfields from their prior publications), even if each of those authors, individually, had published in a less diverse set of subfields. In other words, it is the differences between the background of each co-author that increases MI and not individual authors with diverse backgrounds. Nevertheless, authors having diverse backgrounds may be more likely to increase the MI of one paper, but only if this diversity is sufficiently different from the subfields of the remaining authors. As a result of this approach, a single-author publication, no matter

[10] Porter, A., & Rafols, I. (2012). Interdisciplinarity: Its bibliometric evaluation and its influence in research outputs, 20, p. 21. Retrieved from http://sites.nationalacademies.org/DBASSE/cs/groups/dbassesite/documents/webpage/dbasse_072694.pdf.

[11] Campbell, D. et al. (2015). Application of an "interdisciplinarity" metric at the paper level and its use in a comparative analysis of the most publishing ERA and non-ERA universities. *20th International Conference on Science and Technology Indicators*. Retrieved from http://science-metrix.com/sites/default/files/science-metrix/publications/campbell_et_al_sti2015_short_paper_final_web.pdf.

the diversity of its author's background, will always receive the minimum score, since the indicator is intended to capture diversity across different authors.

4.3.2.4 Highly multidisciplinary papers

In this study, the 10% most highly multidisciplinary papers ($HMP_{10\%}$) indicator was employed as a complementary indicator to MI. It is based on the MI indicator, reflecting the share of papers for a given entity that lies among the 10% most multidisciplinary papers in the respective subfield and year. It reduces the potential effect of outliers in MI.

4.3.3 Network analysis

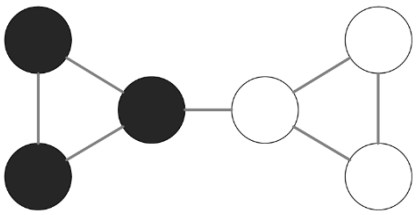
Scientific research is a communal undertaking: collaboration is an important conduit for introducing new perspectives into research, highlighting assumptions, outlining new hypotheses, and sharing testing and analysis methods. The dynamics of this interconnected research ecosystem can be analyzed at the level of individual researchers and institutions, as well as at the network level, where one can find emergent properties that explain features of the ecosystem.

Network-level analyses require a delineation of actors (nodes) and the types of connections between them (edges). With nodes and edges defined, numerous social network analysis tools can be applied to discover the dynamics of the system. Science-Metrix has a wealth of experience working with these tools. Below is a summary of possible network indicators, some of which might be used in this project; preliminary analyses are recommended to demonstrate the potential value of the following:

- Degree, betweenness centrality and eigenvector centrality. These indicators characterize the network properties of individual nodes therein.
- Average degree, heterophily and homophily. These indicators characterize properties of the network as a whole.

Heterophily-homophily indicators were be of particular importance to this study, and the method used to compute these indicators is exposed in greater detail below.

Homophily: Homophily is the tendency of individuals to associate and collaborate with similar others. For example, if researchers collaborate more with other researchers from their country than with researchers from other countries then the network is homophilic in regard to the country. Homophily is thus the opposite of integration. Traditionally, homophily is calculated by comparing the average path length of actors that share a characteristic to the average path length of all actors in the network. In the following example, the average path length between any two actors is 1.8, but the average path length between white actors (or between black actors) is 1, meaning that the network is homophilic in regard to the colour of actors.



However, in the case of networks with disconnected components, this definition of homophily is mostly useless because it becomes impossible to compute the average path length of disconnected actors. This is a problem for this project since many of the networks at stake involve high numbers of disconnected actors. Consequently, a modified version of this indicator had to be developed to provide insights on homophily in the networks. For a given network, the frequency of collaborations for each type of pairs (e.g. pairs of themes, sectors, countries, etc.) is computed based on the total number of collaborations in the network. Then, the number of collaborations of each pair in a random network of the same size (e.g. same number of actors and collaborations) is calculated to obtain a reference level. For homophilic pairs, the expected frequency is:

$$Frequency_{Hom} = \left(\frac{N_1}{N}\right)\left(\frac{N_1 - 1}{N - 1}\right)$$

where N is the total number of actors in the network and N_1 is the number of actors for this specific network level (e.g., sector, thematic, country). For heterophilic pairs, the expected frequency is:

$$Frequency_{Het} = \left(\frac{N_1}{N}\right)\left(\frac{N_2}{N - 1}\right) + \left(\frac{N_2}{N}\right)\left(\frac{N_1}{N - 1}\right)$$

where N is again the total number of actors in the network, and N_1 and N_2 are the numbers of actors for the two specific network levels involved in the pairing (e.g., sectors, thematics, countries). For both homophilic and heterophilic pairs, the expected occurrences are then obtained by multiplying the frequency by the total number of collaborations in the network, which represents the weighted sum across all the dyads in the network.

Next, the sum of observed collaborations for all homophilic pairs is divided by the sum of expected collaborations of all homophilic pairs to provide a ratio of observed over expected frequencies. The same process is repeated for heterophilic pairs, then the ratio of homophilic pairs is divided by the ratio of heterophilic pairs, with a score above 1.00 indicating that homophilic pairs are more present then expected compared to heterophilic pairs and thus the network is homophilic while a score below 1.00 indicates that heterophilic collaborations are more frequent than expected and thus the network is heterophilic. The logarithm value of this score is presented in this report in order to bound the indicator between -1.00 and 1.00, with 0 acting as the neutral value at which the network is neither homophilic nor heterophilic, while scores above 0 indicate homophily and score below 0 indicate heterophily in the network.

While the above indicator presents a homophily score for the network as a whole, ratios of observed over expected frequencies for each type of pairs (e.g. pairs between sectors, thematics, countries, etc.) can be computed in the same manner, providing information on homophily at these levels. For instance, while a network can be homophilic, some pairings might not and thus scores at the pair level helps highlighting different patterns in the network. Ratios over 1.00 indicates pairs which are more frequent than expected in the network while scores below 1.00 indicates the reverse case.

IDRC · CRDI
International Development Research Centre
Centre de recherches pour le développement international

3 November 2021

This letter briefly reflects on IDRC's motivations in commissioning the attached report and what we learned from it. The report is part of an exercise to scope future research investment in climate adaptation and resilience. We asked Science-Metrix to examine how this field has been reshaped (or not) by multi-country and cross-disciplinary research practices in recent years. Specifically, we wanted to see whether these collaborative research practices might have contributed to changes in southern and women authors' network positions.

The two variables of *heterophily* and *centrality* reveal interesting trends over time regarding female authors and those based in global south.

- **Women** accounted for one-third of authorships in recent period, representing only a slight increase since 2005. There remains a persistent gender gap in contributing to scholarship which warrants continued and renewed attention from research funders. This gender gap extends beyond simply contributing to publications, as the report reveals women to be less prominent (or central) within existing networks of research collaboration, although this has improved over time. Women authors appear to be relatively better positioned within the subset of publications that are highly interdisciplinary.

- There is a widening gap between authors based in the **global south** and global north. The prevalence of north-south and south-south co-authorship increased slightly over 2005-19. The share of international co-publications in the field of climate adaptation and resilience rose between 2005-19: from 33% to 38% overall, and from 15% to 26% in the "Topic Cluster 438". Yet northern-based authors remained more central in networks. The report suggests that North-South co-publications may decrease the prominence of southern authors with research networks. The report also suggests that while southern authors increasingly publish as part of north-south collaboration, there is a worrying trend of reduced centrality of authors based in lower- and lower-middle income countries. Open Access was not found to necessarily increasing the connectivity (or centrality) of southern authors within research networks.

In short, gender and geography are distinct forms of **inequality in research networks**. This report suggests that North-south co-publication did not impact the position of women authors. To better understand the intersection of gender and geography, future work could probe the factors that encourage or discourage women to enter into international research collaborations. While the report cannot explain why the above trends arise, it does inspire additional questions. For example, how does collaboration grow over time, how do authors join the research networks, and how can networked be encouraged towards greater equality in participation? Science-Metrix's findings provide evidence in favour of funders being proactive in supporting southern-led research and women researchers on teams.

IDRC is intrigued by the potential for applying **network analysis**. The present work examines the network of authorships, yet also of interest are the networks of topics or ideas reflected in the

HEAD OFFICE / SIÈGE: 150 Kent Street / 150, rue Kent ○ PO Box / CP 8500 Ottawa ON ○ Canada K1G 3H9
Phone / Téléphone : +1 613 696 2176 ○ Email / Courriel : bcurrie-alder@idrc.ca / bcurrie-alder@crdi.ca
idrc.ca / crdi.ca

Canada

scholarship[1]. This could potentially help identify how thinking in the field has evolved over time, or to locate "white spaces" or adjacent opportunities for future investment. The present report identifies Scopus topics 1567 and 3646 as pertinent to climate adaptation & resilience. Reflecting on this report, we learned that IDRC does not necessarily seek to enhance the centrality of funded authors and their work. Certain features are less visible in these networks and suggest how networks might evolve over time, such as: opportunities for early-career researchers, nurturing fresh ideas that are not yet established in topics of prominence, and how real-world utilization of research leads to improved climate resilience. We also acknowledge the limitations of bibliometric databases within which non-English language and non-peer-reviewed outputs are less visible. This can be due to available research being shared in local languages or unindexed, or simply due to lack of research in some places.

The report attempted to situate IDRC support by examining how funded publications and grantees "map" onto the larger field of climate adaptation. It is encouraging that IDRC-FCDO funded research appears to be highly cross-disciplinary, as "30% of these publications were found to be amongst the top decile of publications for multidisciplinarity in [2010-19]". Yet we note that efforts to map funder support against the research field could focus on small portion of overall literature (less than 385,000 publications) and a more robust sample of funded outputs (beyond the ~400 publications provided). We further note Science-Metrix found additional IDRC and FCDO-funded publications by scanning the acknowledgements section of articles, beyond those recorded in our information systems. This may be due to delay in time, as publications emerging after project closure, and suggests an opportunity to look at the publications from past investments beyond the timeframes of our current tracking systems.

IDRC intends to use this report to provoke dialogue with colleagues, grantees and partners to further explore these issues. We acknowledge the need to increase the participation of women researchers and those from developing countries. For example, recent work found that: women and researchers from the global south are under-represented in 100 most highly-cited papers in climate-science[2]; while less than 4% of global research funding for climate change is spent on African topics, and the majority of that funding is channeled to institutions outside the continent.[3] We are committed to understanding the implications of our research funding, and seeking ways to foster a more equitable and diverse climate research landscape.

We express our appreciation for the patience, persistence, and hard work of Etienne Vignola-Gagné, David Campbell, Henrique Pinheiro and colleagues at Science-Metrix. We also thank our colleagues Heidi Braun and Matthew Wallace for their helpful advice and comments.

Sincerely,
Bruce Currie-Alder and Michele Leone

[1] Notable: Baggio (2021) 30 Years of adaptation, mitigation and transformation related to climate change. Climatic Change 10.1007/s10584-021-03146-5; Nalau and Verrall (2021) Mapping the evolution and current trends in climate change adaptation science. Climate Risk Management 10.1016/j.crm.2021.100290; and Lesnikowski et al (2019) Frontiers in data analytics for adaptation research. WIREs Climate Change 10.1002/wcc.576
[2] Tandon (2021) Lack of diversity in climate-science research. Carbon Brief
[3] Overland et al (2021) Funding flows for climate change research on Africa. Climate and Development 10.1080/17565529.2021.1976609